



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 155849

TO: Medina A Ibrahim
Location: REM-2B07/2C18
Art Unit: 1638
Thursday, June 09, 2005

Case Serial Number: 10/718952

From: Toby Port

Location: Biotech-Chem Library
REM1-A59

Phone: 272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Ibrahim,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

This Page Blank (uspto)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop-off or send completed forms to STIC Biotech-Chem Library, Remsen Bldg.



This Page blank (uspro)
BEST AVAILABLE COPY

This Page blank (uspro)


```
Db 241 AGNDTMENTNLASVDKNEAEVSPSTLYAIACWMEGVPIGNSPQMTFVPGILDLAIKNNC 300
Oy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Oy 361 NVVDDMVSNSAILYEPGEHPDHVVVIKVPYVGSKRAMEYTSSEIFMGKNTVLHNTC 420
Db 361 NVVDDMVSNSAILYEPGEHPDHVVVIKVPYVGSKRAMEYTSSEIFMGKNTVLHNTC 420
Oy 421 EDLSLAAPITIDLVLLAEISTRIOFKANEKGKFSFHPVATITLSYTKAPLVPPGTVPVN 480
Db 421 EDLSLAAPITIDLVLLAEISTRIOFKANEKGKFSFHPVATITLSYTKAPLVPPGTVPVN 480
Oy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALAKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 2

```
US-09-677-064-11
; Sequence 11, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Calc, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phycate Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-677-064-11
```

```
Query Match 89.2%; Score 2346; DB 3; Length 510;
Best Local Similarity 87.8%; Pred. No. 2e-222;
Matches 448; Conservative 28; Mismatches 34; Indels 0; Gaps 0;
```

```
Oy 1 MFIEFKVESPNVYKTEETELIOSVYNYETTELVHENRNGTQVIKPKSVNQFTKNTHP 60
Db 1 MFIESFRVESHVRYGPEIESEYRYDTTELVEHKGDSARVVVPKSVKTNFRITRAVP 60
Oy 61 KLGVWLVMGNGNSTLTGVIANREDISWATKDIQOANYGSLTQASAIRVSGFOEE 120
Db 61 KLGVWLVMGNGNSTLTGVIANREDISWATKDIQOANYGSLTQASAIRVSGFOEE 120
Oy 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESWVPLPGI 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESWVPLPGI 180
Oy 181 YDPPIFANQOSRRANNVIKGTKEQVOQI IKDIAFKFATKVDKVVVLTANTERYSNLV 240
Db 181 YDPPIFANQOSRRANSVIKGTKEQVEQI IKDIREFKKNKVDKIVLMTANTERYSNVC 240
Oy 241 VGLNDTMENTLAAVDRNEAEISPSSTLYAIACWMEVNPFIINGSPOMTFVPGILDLAIANT 300
Db 241 VGLNDTMENTLAAVDKNEAEVSPSTLYAIACWMEGVPIGNSPQMTFVPGILDLAIKNNC 300
```

```
Oy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Oy 361 NVVDDMVSNSAILYEPGEHPDHVVVIKVPYVGSKRAMEYTSSEIFMGKNTVLHNTC 420
Db 361 NVVDDMVSNSAILYEPGEHPDHVVVIKVPYVGSKRAMEYTSSEIFMGKNTVLHNTC 420
Oy 421 EDLSLAAPITIDLVLLAEISTRIOFKANEKGKFSFHPVATITLSYTKAPLVPPGTVPVN 480
Db 421 EDLSLAAPITIDLVLLAEISTRIOFKANEKGKFSFHPVATITLSYTKAPLVPPGTVPVN 480
Oy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALAKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 3

```
US-09-727-628-2
; Sequence 2, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727,628
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,612
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-727-628-2
```

```
Query Match 88.8%; Score 2337; DB 4; Length 510;
Best Local Similarity 87.8%; Pred. No. 1.5e-221;
Matches 448; Conservative 28; Mismatches 34; Indels 0; Gaps 0;
```

```
Oy 1 MFIEFKVESPNVYKTEETELIOSVYNYETTELVHENRNGTQVIKPKSVNQFTKNTHP 60
Db 1 MFIESFRVESHVRYGPEIESEYRYDTTELVEHKGDSARVVVPKSVKTNFRITRAVP 60
Oy 61 KLGVWLVMGNGNSTLTGVIANREDISWATKDIQOANYGSLTQASAIRVSGFOEE 120
Db 61 KLGVWLVMGNGNSTLTGVIANREDISWATKDIQOANYGSLTQASAIRVSGFOEE 120
Oy 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESWVPLPGI 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESWVPLPGI 180
Oy 181 YDPPIFANQOSRRANNVIKGTKEQVOQI IKDIAFKFATKVDKVVVLTANTERYSNLV 240
Db 181 YDPPIFANQOSRRANNVIKGTKEQVEQI IKDIREFKKNKVDKIVLMTANTERYSNVC 240
Oy 241 VGLNDTMENTLAAVDRNEAEISPSSTLYAIACWMEVNPFIINGSPOMTFVPGILDLAIANT 300
Db 241 VGLNDTMENTLAAVDKNEAEISPSSTLYAIACWMEGVPIGNSPQMTFVPGILDLAIKNNC 300
Oy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Oy 361 NVVDDMVSNSAILYEPGEHPDHVVVIKVPYVGSKRAMEYTSSEIFMGKNTVLHNTC 420
Db 361 NVVDDMVSNSAILYEPGEHPDHVVVIKVPYVGSKRAMEYTSSEIFMGKNTVLHNTC 420
```

Qy 421 EDSLLAAILDLVLLAELSTRIOFKAENEGFHSFHPVATILSYLTKAPLVPFGTVPVN 480
Db 421 EDSLLAAILDLVLLAELSTRIOFKAENEGFHSFHPVATILSYLTKAPLVPFGTVPVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALAKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 4

US-09-734-237B-73
; Sequence 73, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-73

Query Match 52.8%; Score 1389.5; DB 4; Length 533;
Best Local Similarity 53.1%; Pred. No. 4.9e-128;
Matches 278; Conservative 87; Mismatches 138; Indels 21; Gaps 8;
Qy 3 IENFVESPVNYKYETEISQVYVNYETTELHNNRGTQWIVKPKSVNYQFKTNTHVP-K 61
Db 9 ITSXVAVVDKCTYKQNELLTYSYENA-VVTKTASGRD-VTPVQDVYFKLDLKKREK 65
Qy 62 LGVNLVWGNGNGSTLTGVIANREDISWATKDKIQANFYGSLTQASAIRVG-SFGQEE 120
Db 66 LGIMLIGGNGGSTLVASVLANKHNEVFOKEGVKQPNYFGSMTCSTLKGIDAEQND 125
Qy 121 IYAPKSLIPVNPDDIYFGGWDISNNMLADAMARAKPFDIDLOQLRPYMSMPLPGI 180
Db 126 VYAPNSLIPVNSPDPFVSGWDINNADLYEAMORSQVLEVDLQQLAKKSLVPLPSI 185
Qy 181 YDPDFIAANOBERANNVI-----KGT-----KQEOVOQIIKDIKAKREATKYDKVVLWMTA 231
Db 186 YDPDFIAANOBERANNVI-----KGT-----KQEOVOQIIKDIKAKREATKYDKVVLWMTA 245
Qy 232 NTERYSNLVGLNDTMENTLLAADRNEAISPSTLYAIACWENVPFINGSPONTFVGL 291
Db 246 NTERVVEVSPGVNDTMENTLLQSIKNDHEIASTIFAASILEGVPIYNGSPQNTFVGL 305
Qy 292 IDLAIRNTLLIGDDPFKSGQTRMSVLDVFLVGAIGIKPTISVYNHLGNNDGMUSAPQT 351
Db 306 VOLAEHEGTFLAGDDLKSGQTKLSVLAQFLVDAGIKPVSIASYNHLGNNDGMUSAPQ 365
Qy 352 FRSEKISKNVVDVWNSNAILYEP--GEHPDVVVIKYVVPVSGSKAMDEYSEIFMG 409
Db 366 FRSEKISKNVVDVWNSNAILYEP--GEHPDVVVIKYVVPVSGSKAMDEYSEIFMG 425
Qy 410 GKNITVLNTECEDSLAAILDLVLLAELSTRIOF-----AENEGFHSFHPVATILS 464
Db 426 GHNRSISHNVCEDSLAAILDLVLLAELSTRIOF-----AENEGFHSFHPVATILS 480
Qy 465 YLTAKPLVPFGTVPVYNALSKORAMLENIMRACVGLAPENNMILE 508
Db 486 YMLKAPLTRPGFHPVNGINKORTALENLRLLIGLPSQNEIRFE 529

RESULT 5

US-09-734-237B-75
; Sequence 75, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein derived from Saccharomyces cerevisiae myo-inosi
; OTHER INFORMATION: col-1-phosphate synthase, having a glycine residue inserted after
; OTHER INFORMATION: the initiating methionine
US-09-734-237B-75

Query Match 52.8%; Score 1389.5; DB 4; Length 534;
Best Local Similarity 53.1%; Pred. No. 4.9e-128;
Matches 278; Conservative 87; Mismatches 138; Indels 21; Gaps 8;
Qy 3 IENFVESPVNYKYETEISQVYVNYETTELHNNRGTQWIVKPKSVNYQFKTNTHVP-K 61
Db 10 ITSXVAVVDKCTYKQNELLTYSYENA-VVTKTASGRD-VTPVQDVYFKLDLKKREK 66
Qy 62 LGVNLVWGNGNGSTLTGVIANREDISWATKDKIQANFYGSLTQASAIRVG-SFGQEE 120
Db 67 LGIMLIGGNGGSTLVASVLANKHNEVFOKEGVKQPNYFGSMTCSTLKGIDAEQND 126
Qy 121 IYAPKSLIPVNPDDIYFGGWDISNNMLADAMARAKPFDIDLOQLRPYMSMPLPGI 180
Db 127 VYAPNSLIPVNSPDPFVSGWDINNADLYEAMORSQVLEVDLQQLAKKSLVPLPSI 186
Qy 181 YDPDFIAANOBERANNVI-----KGT-----KQEOVOQIIKDIKAKREATKYDKVVLWMTA 231
Db 187 YDPDFIAANOBERANNVI-----KGT-----KQEOVOQIIKDIKAKREATKYDKVVLWMTA 246
Qy 232 NTERYSNLVGLNDTMENTLLAADRNEAISPSTLYAIACWENVPFINGSPONTFVGL 291
Db 247 NTERVVEVSPGVNDTMENTLLQSIKNDHEIASTIFAASILEGVPIYNGSPQNTFVGL 306
Qy 292 IDLAIRNTLLIGDDPFKSGQTRMSVLDVFLVGAIGIKPTISVYNHLGNNDGMUSAPQT 351
Db 307 VOLAEHEGTFLAGDDLKSGQTKLSVLAQFLVDAGIKPVSIASYNHLGNNDGMUSAPQ 366
Qy 352 FRSEKISKNVVDVWNSNAILYEP--GEHPDVVVIKYVVPVSGSKAMDEYSEIFMG 409
Db 367 FRSEKISKNVVDVWNSNAILYEP--GEHPDVVVIKYVVPVSGSKAMDEYSEIFMG 426
Qy 410 GKNITVLNTECEDSLAAILDLVLLAELSTRIOF-----AENEGFHSFHPVATILS 464
Db 427 GHNRSISHNVCEDSLAAILDLVLLAELSTRIOF-----AENEGFHSFHPVATILS 486
Qy 465 YLTAKPLVPFGTVPVYNALSKORAMLENIMRACVGLAPENNMILE 508
Db 487 YMLKAPLTRPGFHPVNGINKORTALENLRLLIGLPSQNEIRFE 530

RESULT 6

US-09-248-796A-17234
; Sequence 17234, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

```

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17234
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17234

Query Match          52.6%; Score 1384.5; DB 4; Length 525;
Best Local Similarity 52.0%; Pred. No. 1.5e-127;
Matches 266; Conservative 99; Mismatches 132; Indels 15; Gaps 5

QY      9 ESPNVKYETEIQSYVNYETTELVEHNRNGTYQMIVKPKSVNYQFKTTHVPRLGVMVLVG 68
DB      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      13 KSSNSVTIKDDHLTKFTFYENS--VKEKDANG--KFIVTPRASYPEFVKDKVRKVGLLVG 69
DB      :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      69 MGNNGNSTLTGCVIANREDISWATDKIQOANYPFGSLTQASAIRVG--SFOGEIYAPRK 126
DB      |||||||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      70 IGGNGGTTLTGATLADKNHSIFENNEGVAKPNNYGSVTAQSTVKIGVDKETGBDYYVPPN 129
DB      |||||||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      127 SLIPWNRPDIIFYGGWDISNNMADAMARAKFDIDLQQLRPRYMESMYPLPGIYDPDFI 186
DB      ::|||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      130 SILPWNRPNDLVVDMDISGLTLDQAMRAKVLDTYLQQLPYLENKKRPLSIYDPDI 189
DB      |::|||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      187 AANOERRANNVIK-----GTKOEQVOQIIKIDIAKFEATKVDKVVVLMWTANTERYSN 238
DB      190 ALNOSERRANNVNPNQVNGEYKTNNKADVVERIKKIDIEFAKNEIDKVIILMTANTERYAD 249
QY      239 LVVGLNDTMENLLAAVDRREAEISBSTLYAIACWMENVNFINGSPONTPEVPGILDALRAR 298
DB      :::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY      299 NTLIGDDPFKSQOTKMKSVLVDPLVGAQGIKPTISIVSYNHIGNNDGNLSAPOTFRSKETS 358
DB      310 DSFIGDDPFKSQOTKIKSVLQAFLVDAGIKPLPSIASYNHLGNNDGNLSSPQOFRSKETS 369
QY      359 KSNVVDMDVNSNAIILY--EPGEHPDHVVVIKYTPPVYGSKRAMDERTSEIFMGKNTIVL 416
DB      370 KQSVDVDIIESNELLYNKSGSDKHVCITYITKLPAVGDSDKVAMDEYYSMLMGHNKIKSI 429
QY      417 HNTGEDSLAAPILLDVLVLAELSTRIOKAENEKGFSFHRYATLTSLYTAAPLPDGT 476
DB      430 HNVCDSLSLATPLLIDLVVATEFAFATRVQVCKGSKSYDELTPVASLSLXYLKAFIARPGF 489
QY      477 PVNLASKORAMLENIIMRACVGLAEPENNMLE 508
DB      490 KPINGLNKROQLVNLISVLVGLPIDNEIRFE 521

RESULT 7
US-09-902-540-12518
; Sequence 12518, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Goldman, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12518
```

LENGTH: 444;
 TYPE: PRT
 ORGANISM: Myxococcus xanthus
 US-09-902-540-12518
 Query Match 20.1%; Score 528; DB 4; Length 444;
 Best Local Similarity 29.4%; Pred. No. 3.5e-43;
 Matches 135; Conservative 91; Mismatches 187; Indels 46; Gaps 11;
 Db 61 KLGVLVGMGNGNGLTGGVIANREDISWATKDKIQANFSGSLTQASAIRGS-FQGE 119
 13 KLAVALIPGIGANS-TTLMAGVELARKG-----KGPISSTLQMGTAIRGKRDR 61
 QY 120 EYAPFESLLPENVDDIVFGGDISNNQLADAMAPKVPDIDLOKRPYMESWVLP 179
 Db 62 TV--KLNEIVPLAEIKDVAFGAMDIIREDAYEVAVSGVLSDKHLEBVKPLOSIRK 119
 QY 180 IYDPPIANQERANNVKGK--QEQVQIIKDKAFKATKVDKVVVLTMTANTERS 237
 Db 120 VHDPEFVRRIE---NHKATKTHRESTEALRQDIRDFKELNATRAVAVVCSVETFR 175
 QY 238 NLVGLANTMEMLLAADVRENAEISPTLYAICWENVPFINGSPO-NTFVGLDLAI 296
 Db 176 PLPESFK-TLAFERKALDENSPIINTALYTAAIGEPFATATNASVDTALQSMK 234
 QY 297 ANNTLIGDDDFKSGGTQMKSVLVDPLVGAIKPTSIVSNNHGNNDGMNISAPQTF 356
 Db 235 QESVAVAGRDLSQGTMMKVTIAPALKAMGLDGWFSTNIIENRGEVLDDPAAFRAKE 294
 QY 357 ISKSVVVDNMVNSMILVREGHPD-----HYVVVKVYVDSKRAMDEYSEIFMG 410
 Db 295 VTKSSVLDTLL-----QPDLYPELYKKYAHKVALHYPPRQDAEGWMDNITGM 347
 QY 411 KNTIVLHNTCEDSLAARPIIDLVLALSTRIOFAENEGKESHSPVATILSYLTAP 470
 Db 348 PMQIKVNFCLCRSILAPRVLDIALFLDAKLUEWGIQEW-----NSFYKTP 396
 QY 471 LVPGETPVNALSKQPMLENTMRACVGLAPENNMLEY 509
 Db 397 MAQPGIAPVHDLFIQTLKNTLRVAGBDPITHGLDY 435
 RESULT 8
 US-09-248-796A-20645
 ; Sequence 20645, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstein et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 20645
 ; LENGTH: 452
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 US-09-248-796A-20645
 Query Match 4.5%; Score 119.5; DB 4; Length 452;
 Best Local Similarity 21.0%; Pred. No. 0.0071;
 Matches 107; Conservative 80; Mismatches 158; Indels 165; Gaps 31;
 QY 11 PNVKYTER---EIGSVYVE-TTELVA---ENRNGTYQWIVKPKSVNYQPK----- 54
 Db 2 PSIVYPDSTTKQYQATNNSENATEIRYPILESK-----IVNDAENYLLKILQSVIAN 55
 QY 55 ---NTHTVPKGVN-LVGMGNGNGLTGGVIANREDISWATKDKIQANV-FGSLTQA 108

```

Db      56 HPTITNQJPHMLITPRLTWSKROSIEYIKYVLENLEITAFNIIDLSLATPCGVOLTNS 115
Qy      109 SAIRVSGFOGEEIYAPFKSLPMPVNPDDIVFGMDISN-----MNL-----A 150
Db      116 TVVVV---DDENI-----QIVPVVGQALFKFAGKLKNGSITISRELKONLEPNLTSQOI 167
Qy      151 DAMARAKFPD--IDLOKQURPMEBSVPLPGIYDPF-----IAAQEERANNVIKGTQ 203
Db      168 EDLKNSDIFEVAVIIDOCGMALDYIKDITKTNNEDNEFDVAKIVTENONGIPEAIISNTP 227
Qy      204 BOVOQIINDI-KAFPEATK---VDKVV---VLMPTANERYSNLVGMDTMENTENLLAVDRN 257
Db      228 EOQOQEOQDSNPKNELEKNYPIIDSKTOEKIWKG-ERFS---GTN-----N 270
Qy      258 EAEISPTLYALACWENYFPINGSPONFVPLDLAARNTLIGADDFKSGOTKWSV 317
Db      271 LVKLISSTYS--SLSTIPID-KRODCY-----DNIIIVGSIKFTKGLK-EAI 315
Qy      318 LV---DELYV-----GAGI-----KPTSIYSYNHLGNNDGKNTLSAPOTFRS 354
Db      316 LKLNQDYLVEKPNALIDGSIINDPGVNTALIKYQOQSTINDYNEGSGDNNN----- 366
Qy      355 KEISKSNVDDWMSNALIYEPGEHPDHVVVIKYVPYVGDGRANDE----- 401
Db      367 ---SNSN-----SNSNV-----PNSIKLYVYPPYPEWKKPKKXGSGMHDVFLGQ 411
Qy      402 -YTSIEFMG-----GKNTIVLHNTGEDSL 425
Db      412 IYSKOIYSGSSHGGKELFVSGDMYEBEVL 441

```

RESULT 9
 US-09-077-098A-6
 : Sequence 6, Application US/09077098A
 : Patent No. 6544519
 : GENERAL INFORMATION:
 : APPLICANT: TOKUNAGA, Ei-ji
 : SAKAGUCHI, Masaaki
 : MATSUDO, Kazuo
 : HAMADA, Fukuaburo
 : TOKIYOSHI, Sachio
 : TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
 : PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
 : NUMBER OF SEQUENCES: 8
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BROWDY AND NEWMARK
 : STREET: 624 Ninth Street, N.W., Suite 300
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: USA
 : ZIP: 20001
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent in Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/077, 038A
 : FILING DATE: 19-May-1998
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/JP97/03222
 : FILING DATE: 12-SEP-1997
 : APPLICATION NUMBER: JP 27, 148/1996
 : FILING DATE: 19-SEP-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: KORNEAU, Anne M.
 : REGISTRATION NUMBER: 25, 618
 : REFERENCE/DOCKET NUMBER: TOKUNAGA=1
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 202-628-5197
 : TELEFAX: 202-737-3528
 : INFORMATION FOR SEQ ID NO: 6:
 :

```

;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 2042 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: peptide
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-077-098A-6

```

Query Match	4.5%;	Score 118.5;	DB 4;	length 2042;
Best Local Similarity	23.0%;	Pred. No. 0.11;		
Matches 89;	Conservative 53;	Mismatches 146;	Indels 99;	Gaps 21;

```

QY 3 NGYIOUWVAKRSVAYQCKNIHVHPKJLVMLVWGNGNGS:LTVGSVIANRED-----LSMA 91
Db 1160 SATWQW--ANNGVNY--KTNN-----LTVNSQNGTILFG-----MRBDPSVKQITAG 1203
QY 92 TKDKIQOANTFGSL-----TQASATRVGSFOGEETIYAPFKSLPMWNPPDIYF--GCW 142
Db 1204 TYNTTGDPANNKQOAMNTLQOITTEATGITSVSSTYNAGFS-----LGAOSVTFKSGAG 1258
QY 143 DISMNTLADNAPRAKVPDIDLOQLRPYMSWVPLPIYDPDPIAANCEERANNVI----- 198
Db 1259 TVKISGVSDAITADT--DAAITLQVKEKRTTLVG-----DNDDITADRSGGTSIGIYNL 1310
QY 199 ---KGTQEOVOQIHKIDKAFKEATKVDKVVMTANTERYSNLVGLANTMENTL--AAV 254
Db 1311 SLNKGT-----VSATTEKRVSGKIVYEAIRAITGNIFTIGLDDTLINKINPPA 1359
QY 255 DRNEAEISPTLYAIACMENV-----PFINGSPONTFVPGILDIAIARNTLIGG 304
Db 1360 DQDLSNLSBSGKNAITGLVAVVKKTNSPIVEBSPDSBKKKTFVVG--VDFT--DRTTEG 1415
QY 305 D--DFKSGQT-----KMSVLDVLVGAGIKTSTVSTNHLGN--NGCMNISAP 349
Db 1416 DATDDKLLTTSKSVESYVTKNLNFTSDILLSDSGSNATTANDGVKRLSDGFTIKSE 1475
QY 350 Q-TFRSKEIKSNVVDMMVNSNAIIVE 375
Db 1476 NFTLGSKQYNGS-----DSLGMMD 1495

```

RESULT 10
 US-09-107-433-4707
 Sequence 4707, Application US/09107433
 Patent No. 6800744
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
 SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
 THERAPEUTICS
 NUMBER OF SEQUENCES: 5206
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: <Unknown>
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: <Unknown>
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,433
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/ 085131
 FILING DATE: May 12, 1998
 APPLICATION NUMBER: 60/051553
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:

Query Match 4.4%; Score 115; DB 4; Length 842;
Best Local Similarity 24.7%; Pred. No. 0.05;
Matches 97; Conservative 51; Mismatches 134; Indels 110; Gaps 25;

QY 93 KDKIOANVFGSLTQASAIRVGSFQGEELIYAPFKSLPLPVNPDIVFGGMDISNM--NLA 150
DB 460 REGTQRFDFENLRVAGQY-IGSFR---TMTFASIQQYVS-----DLKQVNGGE 505
QY 151 DAMARAKVFD-----IDLQKQLRPYMESWVPLPGIY-----DPDFIANQGEER--A 194
DB 506 GLAGRAAAFDHYQAAGRIISQKSDPF-QSAVLDI-GAYKPISNSTPDAIASEVKNRYAA 563
QY 195 NNVIKQ-----TKQEQVQI IKDIAFKKATKVDKVVYVMTANTERYSLVVGNDT 246
DB 564 QDLVAIGITPPLSKQE--SQVLTLD--AVRSTVNOAI-----SLQGLKRT 608
QY 247 M-ENLLAIVDRNEAISPSTLYAIACV-MENVPFINGS---PONTFV---PGLIDLAIR 298
DB 609 LPPKALRVSASIASGSGPTAYALLLGQDQNOYDNRGIIIPYSGFVSKYKPTLDKYDAK 668
QY 299 NTLIGDDPFKSGQTKMKSVLVDPLVAGIKPTSIYSYNHL-----GNNDGNMISAPQ 350
DB 669 -TVLAGDQMLNPTKAMKD-----AGISAVSIPSDKLTXYTDXKGNAFAYSAQAQ 719
QY 351 -----TRSKKISKSNNV--DDMVNSNAILYEPGE-----HPDHVVYIKYVP 390
DB 720 VAMGNFRSAYALAOQSDASKTNTVSPSDIAEKAAQYATGTVYKGLNGSDVV-----MP 775
QY 391 YVGSKRAMDEYTS---EIFWGGKNTIVLHN 418
DB 776 FGMDKTTFRDRTYTAGREAMTQAGLNPSSLEN 807

RESULT 13

US-09-014-897-2
Sequence 2, Application US/09014897
Patent No. H002085

GENERAL INFORMATION:

APPLICANT: Hoskins, Johann
Jaskunas, S. Richard
Rockey, Pamela K.
Zhao, Genshi
Rosteck, Paul R. Jr.
No. H002085tis, Franklin H.
TITLE OF INVENTION: Penicillin Binding Protein From
Streptococcus Pneumoniae

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,897
FILING DATE: 28-Jan-1998
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/731,716
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-10,887
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 731 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-014-897-2

Query Match 4.3%; Score 114; DB 1; Length 731;
Best Local Similarity 20.8%; Pred. No. 0.05;
Matches 108; Conservative 64; Mismatches 183; Indels 165; Gaps 29;

QY 90 WATDKIOQANVFG-----SLTQASAIRVGSFQGEELIYAPFKSLPLPVNPDIVFGGMD 143
DB 215 WGEVDASKK--YFGVSASEVSLDQAAVL--AGMLKGPGLYNPLNSVEDSTNRDDTV----- 266
QY 144 ISNNMLA---DAMARAKVFDIDLQKQLRPYMESWV---LPGIYDPDFIAA-----NOE 191
DB 267 LQNVYAAGYIDKQGETEALEVDMTSQLDKTEGKISDRYRSPYFPAVVNEAVSKYNLTBE 326
QY 192 ERANN-----VIKGTQEQVQI IKDIAFKEA-----TKVDKVVYVLM 229
DB 327 EIVNNGYRIYTELDQNVQANNQIYVENTSLFPRAEDGTFAQSGSVALBPKTGVRGVVQ 386
QY 230 TANTER-----YSN-----LVY-----GLNDMENTLAAVDRNEAEI 261
DB 387 VADNDKTFRRNFVNATQSRKSPGSTIKPLVYTTPAVEAGMALNKQLDHTMQDYDKVD- 445
QY 262 SPSTLYAIACWENVPFINGSFONTFVPLI---DLAIRNTLIGDDPFKSGQTKMKSVL 318
DB 446 ---NYAGIKTSREVPMTQSLAESLNLPAVATVNDLVGDK-AFEAGEKFGGLMEVYDKVL 500
QY 319 VDFLVGAGIK--PISIVSYNHLGNNDGN-----NLSAPQTFRSKEISKSNNVVD 365
DB 501 -GVMLGSGVETNPLOMAQAYAAAFANBGLMPEAHFISRIENMSG-QVIAISHKNSQKRVLDK 558
QY 366 MV-----NSNAILYEPGEHPDHVVYIKYVYVGSKRAMD-EITSEIFPMG- 410
DB 559 SVADKMTSMMLGTFTNGTGISSPA---DYWAGK---TGTEVVFPEYTSDDWVGIV 611
QY 411 -KNITVLH-----NTCEDSLAAPTIIIDVLVLAELSTRIOFKAENEGKFSHPVA-TIL 463
DB 612 TPDVVISHWLGPPTDENHYLAG-----STNSGAHVFRNIANTITL 652
QY 464 SYLTKAPLVPPGTPVNVNLSKQRALEINIMRACVGLADEN 503
DB 653 -----PYTPGSTFTVENAYAKON-----GIADAN 675

RESULT 14

US-08-731-716-2
Sequence 2, Application US/08731716
Patent No. 5789202

GENERAL INFORMATION:

APPLICANT: Hoskins, Johann
Jaskunas, S. Richard
Rockey, Pamela K.
Zhao, Genshi
Rosteck, Paul R. Jr.
No. 5789202tis, Franklin H.
TITLE OF INVENTION: Penicillin Binding Protein From
Streptococcus Pneumoniae

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```
?
?      COMPUTER: IBM PC compatible
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: PatentIn Release #1.0, Version #1.30
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER: US/08/731,716
?      FILING DATE:
?      CLASSIFICATION: 435
?      ATTORNEY/AGENT INFORMATION:
?      NAME: Webster, Thomas D.
?      REGISTRATION NUMBER: 39, 872
?      REFERENCE/DOCKET NUMBER: X-10, 887
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 317-276-3334
?      INFORMATION FOR SEO ID NO: 2:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 731 amino acids
?      TYPE: amino acid
?      TOPOLOGY: linear
?      MOLECULE TYPE: protein
?      IS-08-731-716-2
```

Query Match	4.3%;	Score 114;	DB 1;	Length 731;
Best Local Similarity	20.8%;	Pred. No. 0.055;		
Matches 108;	Conservative 64;	Mismatches 183;	Indels 165;	Gaps 29

[illegible]

RESULT 15
 US-09-489-039A-8455
 ; Sequence 8455, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2/09,2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27

: PRIOR APPLICATION NUMBER: US 60/117,747
 : PRIOR FILING DATE: 1999-01-29
 : NUMBER OF SEQ ID NOS: 14342
 : SEQ ID NO 8455
 : LENGTH: 778
 : TYPE: PR1
 : ORGANISM: Klebsiella pneumoniae
 : US-09-489-039A-8455

Query Match	4.3%;	Score 113.5;	DB 4;	Length 778;
Best Local Similarity	19.3%;	Pred. No. 0.068;		
Matches 78;	Conservative 61;	Mismatches 149;	Indels 117;	Gaps 18;

QY 68 GWCNNNGSLTGTGVIANREDISMAKDXIQO-----ANVFSLTQASAIRV-----1133
 Db 160 GMRGDDSDS-----KGDTSWVAEMIDHIDVIKRGAAARVNGAMGGVNVIVTKPTT 21.10
 QY 114 ---GSF-----QGEETVAFKSLPMPVP--DDIVFGGMDISNNMLADAM-----1533
 Db 211 PEWHGSMNTYNNAPDHRKKGATKRTNFSLGNPLSDSVSFNLWGNLSKTOADQDINAGHE 27.07
 QY 154 -----ARAKYFDIDLQOLARPYMESWPLPGIYDPPFIANADEFKANNVIK 20.00
 Db 271 AERTGSYAGSYPAGEGVNNDIHSKLMFEPAPMOAL-----EFEG--YSRGGNLYAG 3.22
 QY 201 -TKQEGVQOIIDIAFKKATKVDKVV--VLTWNTATE-----RYNVLVGL 24.43
 Db 323 DTIONTTSTLVSMTWG-KETNRLIYQTYGVMTGMDNGVTSNSYAQYEHTRNSRMBGL 38.81
 QY 244 NDTMENILAAVRNEAIESPTLYAIAACMENVPFINGSPONTEVPGILDAIARNTLIG 30.33
 Db 382 AGGTBEIIFSSSEFSDIDLADVLHSEV-----NIPFLGVQDN-----LTL 42.22
 QY 304 GDDFKSGQTKKSVLVDPLVGAQIKPISIVSYNHLGNNDGNLSAPOTFKSKISKNV 36.33
 Db 423 GTEW--NQGRMD-----GVSTTQALSY--GLTDGVSAAGRSPYSASEIFSLFTE 46.68
 QY 364 DDMVNSNALIVERGHPDH--VVYIKYPPYVGDSKRAMDEVTSI 40.6
 Db 469 DNNALTDSTMLPALRFHSHIVGNMSPSANTLSQELTDWTLTL 51.3

Search completed: June 7, 2005, 16:48:55
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 15:09:19 ; Search time 92.5 Seconds
(without alignments)
2132.409 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632

Sequence: 1 MEIKPKVCEPVMKYTEETI.....NIMRACVGLAPENNILEYK 510

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2632	100.0	510	2	AAW79740 Soybean w
2	2632	100.0	510	8	ADQ14449 Mutant so
3	2632	100.0	510	8	ADQ14491 Wild type
4	2632	100.0	510	8	ADSB2000 Soybean m
5	2632	100.0	510	8	ADSB1994 Soybean m
6	2637	99.8	510	2	AAW79741 Soybean m
7	2637	99.8	510	8	ADQ14495 Mutant so
8	2637	99.8	510	8	ADSB1998 Soybean m
9	2600	98.8	510	8	ADQ14503 Mutant so
10	2600	98.8	510	8	ADQ14505 Wild type
11	2600	98.8	510	8	ADSB2004 Soybean m
12	2600	98.8	510	8	ADSB2006 Soybean m
13	2593	98.5	510	8	ADQ14501 Mutant so
14	2593	98.5	510	8	ADSB2002 Soybean m
15	2472	93.9	536	2	AA124477 Nicotiana
16	2432	92.4	505	8	ADG73738 Myo-inos
17	2378	90.3	510	8	AAAG50575 Arabidops
18	2378	90.3	510	8	ADN73525 Thale cre
19	2378	90.3	581	3	AAAG50574 Arabidops
20	2378	90.3	645	3	AAAG50573 Arabidops
21	2368	90.0	510	4	AAAB8935 Brassica
22	2352.5	89.4	511	3	AAAG9861 Arabidops
23	2352.5	89.4	511	3	AAAG32501 Arabidops
24	2352.5	89.4	534	3	AAAG9860 Arabidops
25	2344	89.1	510	2	AAW96259 Phytate p

ALIGNMENTS

26	2340	88.9	510	8	ADP43920	Adp43920 Rice L-my
27	2336	88.8	510	2	AAW97882	AAW97882 Maize myo
28	2335	88.7	510	4	AAAB47286	AAAB47286 MIP synth
29	2080	79.0	446	3	AAAG32502	AAAG32502 Arabidops
30	2080	79.0	446	3	AAAG09862	AAAG09862 Arabidops
31	1871	71.1	512	8	ADP43919	ADP43919 Porterea
32	1768	67.2	380	3	AAAG32503	AAAG32503 Arabidops
33	1607	61.1	558	4	AAAG5577	AAAG5577 Human hMT
34	1607	61.1	558	4	AAAB40076	AAAB40076 Human pol
35	1607	61.1	558	4	AAAB92849	AAAB92849 Human pro
36	1607	61.1	558	4	AAAB93732	AAAB93732 Human pro
37	1607	61.1	558	5	AAO18891	AAO18891 Human ova
38	1596.5	60.7	534	6	ABJ25857	ABJ25857 Aspergill
39	1592.5	60.5	557	4	AAAG5576	AAAG5576 Murine mT
40	1592.5	60.5	534	6	ABJ26457	ABJ26457 Aspergill
41	1585	60.2	565	4	ABAB5497	ABAB5497 Drosophi
42	1556.5	59.1	659	8	ADN99714	ADN99714 Novel hum
43	1525	57.9	505	8	ADG73737	ADG73737 Aspergill
44	1525	57.9	532	8	ADG73736	ADG73736 Aspergill
45	1420	54.0	504	7	ADB64735	ADB64735 Human pro

RESULT 1
ID AAW79740 standard; protein, 510 AA.
XX AAW79740;
XX

AC 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
DT

DE Soybean wild-type myo-inositol 1-phosphate synthase.
DE

XX Soybean, myo-inositol 1-phosphate synthase; raffinose; stachyose;
XX phytic acid.
XX

XX Glycine max; line LR13.
XX

XX WO9845448-A1.
XX

XX 15-OCT-1998.
XX

XX 07-APR-1998; 98WO-US006822.
XX

XX 08-APR-1997; 97US-00835751.
XX

XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX

XX Hitz WD, Sebastian SA;
XX

XX MPI; 1998-568353/48.
XX

XX N-PSDB; AAW62440.
XX

PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PI phytic acid, etc.
PI

PS Example 5; Page 45-47; 63pp; English.
PS

CC This is the amino acid sequence of soybean myo-inositol 1-phosphate
CC synthase (MI 1-PS) deduced from the coding region of an isolated cDNA
CC clone (see AAW62440) . MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. A mutant MI 1-PS (see AAW79741) has been
CC identified in soybean line LR33, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC in the LR33 gene sequence that resulted in a K396N substitution in the
CC mutant protein. The mutation results in a seed phenotype of very low
CC raffinose saccharide sugars, very high sucrose and low phytic acid. The
CC mutated nucleic acid is used to alter the raffinose saccharide, sucrose,
CC phytic acid and inorganic phosphate content of soybean seeds, leading to

CC useful soybean products, e.g. a seed phytic acid content of less than 17
 CC ug/g, a seed content of raffinose and stachyose combined of less than
 CC 14.5 ug/g, and a seed sucrose content greater than 200 ug/g. (Updated on
 CC 17-OCT-2003 to standardise OS field)

XX Sequence 510 AA;

Query Match 100.0%; Score 2632; DB 2; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1,2e-212;
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MEIENPKVECPNVKKTETETISGVNYVETTELVEHNRNGTYQWIVPKSVKYEFTKNIHVP 60
DB 1 MEIENPKVECPNVKKTETETISGVNYVETTELVEHNRNGTYQWIVPKSVKYEFTKNIHVP 60
QY 61 KLGWLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
DB 61 KLGWLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
QY 121 IYAPFKSLIPWNPDDIVFGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDDIVFGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVIKGTQEQVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVIKGTQEQVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240
QY 241 VGLNDTMENLLAADVDRNEAISPSTLYAIACMENVPIFGSPQTFVGLIDLAIAANT 300
DB 241 VGLNDTMENLLAADVDRNEAISPSTLYAIACMENVPIFGSPQTFVGLIDLAIAANT 300
QY 301 LIGGDDFKSGQTKMSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSIEIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSIEIFMGKSTIVLHNTC 420
QY 421 EESLLAAPTIIIDLVNLAELSTRIEPKAENEGKFSHPATILSYITKALVPFGPPVN 480
DB 421 EESLLAAPTIIIDLVNLAELSTRIEPKAENEGKFSHPATILSYITKALVPFGPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 2

ADQ14499
 ID ADQ14499 standard; protein; 510 AA.

XX ADQ14499;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #2.

XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
 KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
 mutant; mutain.

XX Glycine max.
 OS Synthetic.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00293315.
 PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI: 2004-53135/51.

DR N-PSDB; ADQ14498.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 useful for producing plants with decreased raffinose, stachyose, and
 phytic acid and increased sucrose, leading to valuable and useful soybean
 products.

PS Example 8; SEQ ID NO 10; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant,
 CC a seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents a mutant soybean myo-inositol
 CC 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 100.0%; Score 2632; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1,2e-212;
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MEIENPKVECPNVKKTETETISGVNYVETTELVEHNRNGTYQWIVPKSVKYEFTKNIHVP 60
DB 1 MEIENPKVECPNVKKTETETISGVNYVETTELVEHNRNGTYQWIVPKSVKYEFTKNIHVP 60
QY 61 KLGWLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
DB 61 KLGWLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
QY 61 KLGWLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
DB 61 KLGWLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
QY 121 IYAPFKSLIPWNPDDIVFGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDDIVFGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 121 IYAPFKSLIPWNPDDIVFGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDDIVFGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVIKGTQEQVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVIKGTQEQVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240
QY 181 YDPDFIAANOEBRANNVIKGTQEQVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVIKGTQEQVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240
QY 241 VGLNDTMENLLAADVDRNEAISPSTLYAIACMENVPIFGSPQTFVGLIDLAIAANT 300
DB 241 VGLNDTMENLLAADVDRNEAISPSTLYAIACMENVPIFGSPQTFVGLIDLAIAANT 300
QY 301 LIGGDDFKSGQTKMSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
QY 301 LIGGDDFKSGQTKMSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSIEIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSIEIFMGKSTIVLHNTC 420

```

Oy	421	EDSLAAPPIILVLVLAELSTRIIEPAENKGRKSPHVAITISLYLTKAPLVPPGPPVYN	480
Db	421	EDSLAAPPIILDLVLVAELSTRIIEPAENKGRKSPHVAITISLYLTKAPLVPPGPPVYN	480
Oy	481	ALSKORAMLEINIMRACVGLAPENNMILEYK	510
Db	481	ALSKORAMLEINIMRACVGLAPENNMILEYK	510
RESULT 3			
ID	ADQ14491		
AC	ADQ14491	standard; protein; 510 AA.	
XX	ADQ14491;		
DT	23-SEP-2004	(first entry)	
DE	Wild type soybean myo-inositol 1-phosphate synthase polypeptide #1.		
KW	Soybean: myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;		
KW	raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.		
OS	Glycine max.		
PN	US2004128713-A1.		
XX	01-JUL-2004.		
PD	21-NOV-2003; 2003US-00718952.		
PF	08-APR-1997; 97US-00835751.		
PR	07-APR-1998; 98MO-US006822.		
PR	26-APR-1999; 98US-00299315.		
PR	11-MAR-2002; 2002US-00025003.		
PA	(HITZ/) HITZ W D.		
PA	(SEBA/) SEBASTIAN S A.		
PA	(GRAC/) GRACE D J.		
PA	(STRE/) STREIT L G.		
XX	Hitz WD, Sebastian SA, Grace DJ, Streit LG;		
XX	WPI; 2004-533135/51.		
DR	N-PSDB; ADQ14490.		
XX			
PT	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,		
PT	useful for producing plants with decreased raffinose, stachyose, and		
PT	phytic acid and increased sucrose, leading to valuable and useful soybean		
XX	products.		
XX	Claim 3; SEQ ID NO 2; 48pp; English.		
XX			
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-		
CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate		
CC	synthase having a decreased capacity for the synthesis of myo-inositol 1-		
CC	phosphate. The invention also relates to a chimeric gene operably linked		
CC	to suitable regulatory sequences, where expression of the chimeric gene		
CC	results in a decrease in expression of an endogenous or native gene		
CC	encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant		
CC	comprising the chimeric gene, a method of making the soybean plant, a		
CC	seed of the soybean plant, a soy protein product derived from the		
CC	processing of soybean seeds, a method of making or producing a soy		
CC	protein product and a method of using a soybean plant homozygous for at		
CC	least one gene encoding a mutant myo-inositol 1-phosphate synthase having		
CC	decreased capacity for the synthesis of myo-inositol 1-phosphate. The		
CC	nucleic acid fragment and methods are useful for producing plants with		
CC	decreased raffinose, stachyose and phytic acid content and increased		
CC	sucrose and inorganic phosphate content, leading to valuable and useful		
CC	soybean products. This sequence represents a wild type soybean myo-		
CC	inositol 1-phosphate synthase polypeptide of the invention.		
XX	Sequence 510 AA;		

Query Match	100.0%	Score 2633;	DB 8;	Length 510;
Best Local Similarity	100.0%	Pred. No. 1,2e-212;		
Matches 510;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MFENFKECNNVKTETEIOSVYVNETTELVHNRNGTYOMIYKPKSKVKEFTKTHIVP	60	
DB	1	MFENFKECNNVKTETEIOSVYVNETTELVHNRNGTYOMIYKPKSKVKEFTKTHIVP	60	
QY	61	KLGVMLVGMGNGNSTLIGSVIANREGISWATKDKIQOANYFGSLTQASAIRVSGQEE	120	
DB	61	KLGVMLVGMGNGNSTLIGSVIANREGISWATKDKIQOANYFGSLTQASAIRVSGQEE	120	
QY	121	IYAPFKSLPMPVNPDDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI	180	
DB	121	IYAPFKSLPMPVNPDDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI	180	
QY	181	YDPDFIANOEERANNVYKGTKOBOVOGIIKDIAKPKATKVDKVYVLMTANTRYSNLV	240	
DB	181	YDPDFIANOEERANNVYKGTKOBOVOGIIKDIAKPKATKVDKVYVLMTANTRYSNLV	240	
QY	241	VGNDYTMENLAAVDREAEISPSSTLYALACVMENVPFINGSPONTFVPGILDLAIANT	300	
DB	241	VGNDYTMENLAAVDREAEISPSSTLYALACVMENVPFINGSPONTFVPGILDLAIANT	300	
QY	301	LIGGDDFKSGGCTKAKSVLVDFLVAGIKFPTISVYNHLGNNDGNLSPQTFRSKEISKS	360	
DB	301	LIGGDDFKSGGCTKAKSVLVDFLVAGIKFPTISVYNHLGNNDGNLSPQTFRSKEISKS	360	
QY	361	NVVDMDNNSNALVEPGEHPDHVVIVKVPVVGSKRAMDEYTSIEFMGKSTVLNHTC	420	
DB	361	NVVDMDNNSNALVEPGEHPDHVVIVKVPVVGSKRAMDEYTSIEFMGKSTVLNHTC	420	
QY	421	EDSLAAPRIIDLVLAEISTRIEFKAENEGKFHSFHPVATITLSYLTAKPLVPPTPVVN	480	
DB	421	EDSLAAPRIIDLVLAEISTRIEFKAENEGKFHSFHPVATITLSYLTAKPLVPPTPVVN	480	
QY	481	ALSKORAMLENIMRACVGLAPENNMIIEYK 510		
DB	481	ALSKORAMLENIMRACVGLAPENNMIIEYK 510		
RESULT 4				
AD82000	ID	AD82000	standard; protein; 510 AA.	
AC	AD82000;			
XX	AD82000;			
XX	18-NOV-2004	(first entry)		
DT	18-NOV-2004	(first entry)		
XX	18-NOV-2004	(first entry)		
DE	Soybean myo-inositol 1-phosphate synthase wild-type 3.			
XX	Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;			
KM	raffinose; stachyose; sucrose; inorganic phosphate; flatulence.			
XX	Glycine max; line 29004JP01.			
OS	Glycine max; line 29004JP01.			
XX	Glycine max; line 29004JP01.			
PN	US2003074685-A1.			
XX	US2003074685-A1.			
PD	17-APR-2003.			
XX	17-APR-2003.			
PF	11-MAR-2002; 2002US-00025003.			
XX	11-MAR-2002; 2002US-00025003.			
PR	08-APR-1997; 97US-00835751.			
XX	08-APR-1997; 97US-00835751.			
PR	07-APR-1998; 98WO-US006822.			
XX	07-APR-1998; 98WO-US006822.			
PA	(HITZ/) HITZ W D.			
XX	(HITZ/) HITZ W D.			
XX	(SEBA/) SEBASTIAN S A.			
XX	(SEBA/) SEBASTIAN S A.			
PI	Hitz WD, Sebastian SA,			
XX	Hitz WD, Sebastian SA,			
DR	WPI; 2004-639957/62.			
XX	WPI; 2004-639957/62.			
DR	N-PSDB; AD81999.			
XX	N-PSDB; AD81999.			

CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 100.0%; Score 2632; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.2e-212;
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MEIENFKVCECPVKKYETEISQVYNYETTELVEHNRNGTYQIVKPKSVKYEFTKNIHVP 60
DB 1 MEIENFKVCECPVKKYETEISQVYNYETTELVEHNRNGTYQIVKPKSVKYEFTKNIHVP 60
QY 61 KLGWLVGVGNGNGSTLTGCVIYANREGISMAWKDKIQOANYFGSLTQASAIRVGSFGQGE 120
DB 61 KLGWLVGVGNGNGSTLTGCVIYANREGISMAWKDKIQOANYFGSLTQASAIRVGSFGQGE 120
QY 121 IYAPFKSLIPVNPDPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPVNPDPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPDFIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAANDRNEAISPSTLYAACWENVPFINSPONTFVPGILDIAIARNT 300
DB 241 VGLNDTMENLLAANDRNEAISPSTLYAACWENVPFINSPONTFVPGILDIAIARNT 300
QY 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYVNHGNDGMNLSAPOTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYVNHGNDGMNLSAPOTFRSKEISKS 360
QY 361 NVVDPMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDPMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPITIIDLVLLAEISTRIEPAKANEKGKFSFHVATILSYLTKAPLVPGTPPVN 480
DB 421 EDSLLAAPITIIDLVLLAEISTRIEPAKANEKGKFSFHVATILSYLTKAPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 6
AAW79741
ID AAW79741 standard; protein; 510 AA.
XX
AC AAW79741;
XX
XX 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
DE Soybean mutant myo-inositol 1-phosphate synthase.
XX
XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid.
XX
XX Glycine max; line LR33.
OS
XX
XX WO9845448-A1.
XX
XX 15-OCT-1998.
PD
XX
XX 07-APR-1998; 98WO-US006822.
PF
XX
XX 08-APR-1997; 97US-00835751.
PR
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA
XX
XX Hitz WD, Sebastian SA;
PI

```

XX WPI: 1998-568353/48.
 DR N-PSDB; AAW62443.
 XX

PT Soybean plants containing altered myo-inositol 1-phosphate gene - useful
 PT for generating plants with altered levels of e.g. raffinose, stachyose,
 PT phytic acid, etc.
 XX

PS Example 5; Page 49-51; 63pp; English.

CC This is the amino acid sequence of a mutant soybean myo-inositol 1-
 CC phosphate synthase (MI 1-PS) deduced from the coding region of an
 CC isolated cDNA clone (see AAW62443). MI 1-PS is involved in glucose
 CC metabolism to phytic acid, raffinose and stachyose. The MI 1-PS was
 CC identified in soybean line LR33, a mutagenised line of low raffinose
 CC saccharide phenotype. Sequencing revealed a single base change mutation
 CC in the LR33 gene sequence that resulted in a K396N substitution in the
 CC mutant protein compared to wild-type MI 1-PS (see AAW79740). The mutation
 CC results in a seed phenotype of very low raffinose saccharide sugars, very
 CC high sucrose and low phytic acid. The mutated nucleic acid is used to
 CC alter the raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds, leading to useful soybean products,
 CC e.g. a seed phytic acid content of less than 17 ug/g, a seed content of
 CC raffinose and stachyose combined of less than 14.5 ug/g, and a seed
 CC sucrose content greater than 200 ug/g. (Updated on 17-Oct-2003 to
 CC standardise OS field)

XX Sequence 510 AA;

Query Match 99.8%; Score 2627; DB 2; Length 510;
 Best Local Similarity 99.8%; Pred. No. 3.1e-212;
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MEIENFKVCECPVKKYETEISQVYNYETTELVEHNRNGTYQIVKPKSVKYEFTKNIHVP 60
DB 1 MEIENFKVCECPVKKYETEISQVYNYETTELVEHNRNGTYQIVKPKSVKYEFTKNIHVP 60
QY 61 KLGWLVGVGNGNGSTLTGCVIYANREGISMAWKDKIQOANYFGSLTQASAIRVGSFGQGE 120
DB 61 KLGWLVGVGNGNGSTLTGCVIYANREGISMAWKDKIQOANYFGSLTQASAIRVGSFGQGE 120
QY 121 IYAPFKSLIPVNPDPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPVNPDPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPDFIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAANDRNEAISPSTLYAACWENVPFINSPONTFVPGILDIAIARNT 300
DB 241 VGLNDTMENLLAANDRNEAISPSTLYAACWENVPFINSPONTFVPGILDIAIARNT 300
QY 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYVNHGNDGMNLSAPOTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYVNHGNDGMNLSAPOTFRSKEISKS 360
QY 361 NVVDPMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDPMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPITIIDLVLLAEISTRIEPAKANEKGKFSFHVATILSYLTKAPLVPGTPPVN 480
DB 421 EDSLLAAPITIIDLVLLAEISTRIEPAKANEKGKFSFHVATILSYLTKAPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 7
 ADQ14495
 ID ADQ14495 standard; protein; 510 AA.

XX AC ADQ14495;
XX DT 23-SEP-2004 (first entry)
XX DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX KM Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
XX KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
XX KM mutant; mutein.
XX OS Glycine max.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 396 /note= "Wild type Lys substituted by Asn"
XX PD US2004128713-A1.
XX PD 01-JUL-2004.
XX PF 21-NOV-2003; 2003US-00718952.
XX PR 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX PR 26-APR-1999; 99US-00293315.
XX PR 11-MAR-2002; 2002US-00025003.
XX PA (HITZ/) HITZ W D.
XX PA (SEBA/) SEBASTIAN S A.
XX PA (GRAC/) GRACE D J.
XX PA (STRE/) STREIT L G.
XX PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX DR WPI; 2004-53335/51.
XX DR N-PSDB; ADQ14494.
XX PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX PT useful for producing plants with decreased raffinose, stachyose, and
XX PT phytic acid and increased sucrose, leading to valuable and useful soybean
XX PT products.
XX PS Claim 9; SEQ ID NO 6; 48pp; English.
XX CC The invention relates to a nucleic acid fragment encoding a soybean myo-
XX CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX CC phosphate. The invention also relates to a chimeric gene operably linked
XX CC to suitable regulatory sequences, where expression of the chimeric gene
XX CC results in a decrease in expression of an endogenous or native gene
XX CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX CC comprising the chimeric gene, a method of making the soybean plant, a
XX CC seed of the soybean plant, a soy protein product derived from the
XX CC processing of soybean seeds, a method of making or producing a soy
XX CC protein product and a method of using a soybean plant homozygous for at
XX CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX CC nucleic acid fragment and methods are useful for producing plants with
XX CC decreased raffinose, stachyose and phytic acid content and increased
XX CC sucrose and inorganic phosphate content, leading to valuable and useful
XX CC soybean products. This sequence represents a mutant soybean myo-inositol
XX CC 1-phosphate synthase polypeptide of the invention.
XX SQ Sequence 510 AA;
Query Match 99.8%; Score 2627; DB 8; Length 510;
Best Local Similarity 99.8%; Pred. No. 3, 1e-212;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MFIENFKVCECPNVKXTETETIGSVNYETTELVEHNNNGTYQMIIVKPKSVKFEFKTNHVP 60
|||||

DB 1 MFIENFKVCECPNVKXTETETIGSVNYETTELVEHNNNGTYQMIIVKPKSVKFEFKTNHVP 60
QY 61 KLGWMLVGMGNGNSTLITGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFOEE 120
DB 61 KLGWMLVGMGNGNSTLITGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFOEE 120
QY 121 IYAPFKSILPMVNPDDIVFGGWDISNNMLADAMAPAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSILPMVNPDDIVFGGWDISNNMLADAMAPAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIANQEEERANNVKGTKQEOVOQIIKDIAFKKATVVDKVVYLTMTANTERSYSLV 240
DB 181 YDPDFIANQEEERANNVKGTKQEOVOQIIKDIAFKKATVVDKVVYLTMTANTERSYSLV 240
QY 241 VGLNDTMENLLAAVDRNEAEISPSITLYAIACMVENVPEFINGSPONTFVPGILDLAIARNT 300
DB 241 VGLNDTMENLLAAVDRNEAEISPSITLYAIACMVENVPEFINGSPONTFVPGILDLAIARNT 300
QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLAPQTFRSKEISKS 360
QY 361 NVVDDMVNSNALIVPGEHPDHVVVVKVVPVGDSKRAMDEYTSIEIFWGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIVPGEHPDHVVVVKVVPVGDSKRAMDEYTSIEIFWGKSTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIEFKAKENEGKFSFHPVATILSYLTAPLPVPGTPVYN 480
DB 421 EDSLAAPIIIDLVLAELSTRIEFKAKENEGKFSFHPVATILSYLTAPLPVPGTPVYN 480
QY 481 ALSKQRAMLENIMRACVGLAPEENNMIIEYK 510
DB 481 ALSKQRAMLENIMRACVGLAPEENNMIIEYK 510
RESULT 8
ADSB1998
ID ADSB1998 standard; protein, 510 AA.
XX AC ADSB1998;
XX DT 18-NOV-2004 (first entry)
XX DE Soybean myo-inositol 1-phosphate synthase mutant #1.
XX KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
XX KM phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
XX KM flaculence; mutant.
XX OS Glycine max; line LR33.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 396 /note= "Wild-type Lys substituted by Arg"
XX FT US2003074685-A1.
XX PD 17-APR-2003.
XX PF 11-MAR-2002; 2002US-00025003.
XX PR 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX PA (HITZ/) HITZ W D.
XX PA (SEBA/) SEBASTIAN S A.
XX PI Hitz WD, Sebastian SA;
XX DR WPI; 2004-639957/62.
XX DR N-PSDB; ADSB1997.

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 9; SEQ ID NO 6; 34bp; English.

XX
XX
XX
CC The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a mutant myo-inositol 1-phosphate synthase.

CC Sequence 510 AA;

Query Match 99.8%; Score 2627; DB 8; Length 510;

Best Local Similarity 99.8%; Pred. No. 3.1e-212;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPVNKTETETIQSYNNTETELVHENGTYQIVKPKSVKTEFKNIHP 60
DB 1 MEIENFKVCEPVNKTETETIQSYNNTETELVHENGTYQIVKPKSVKTEFKNIHP 60
QY 61 KGVNVLVGVGNGNGSTLTGCVIANRSGISMATFKDITQOANVGSGLTQASAIRVGSFGGE 120
DB 61 KGVNVLVGVGNGNGSTLTGCVIANRSGISMATFKDITQOANVGSGLTQASAIRVGSFGGE 120
QY 121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVDFIDLOKOLRPMESMLPLPGI 180
DB 121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVDFIDLOKOLRPMESMLPLPGI 180
QY 181 YDPDTIANQERANNVKGTQOEVOQIIKDIKAFKATKVDKVVLTANTERTSYNLV 240
DB 181 YDPDTIANQERANNVKGTQOEVOQIIKDIKAFKATKVDKVVLTANTERTSYNLV 240
QY 241 VGLNMTMENTLAAVBRNEAISPSTLYAACWENVPFINGSPONTFVGLDIDLAIRMT 300
DB 241 VGLNMTMENTLAAVBRNEAISPSTLYAACWENVPFINGSPONTFVGLDIDLAIRMT 300
QY 301 LIGGDFKSGQTKMKSVLDFLVGAGIKRPSIVSYNHLGNNDGMNLAPQTRSKSISKS 360
DB 301 LIGGDFKSGQTKMKSVLDFLVGAGIKRPSIVSYNHLGNNDGMNLAPQTRSKSISKS 360
QY 361 NVVDDMVNSNALIYEGEHPDVVVKYPPYVGDSCRANDTTSSTFPMGSKSTIVLANTC 420
DB 361 NVVDDMVNSNALIYEGEHPDVVVKYPPYVGDSCRANDTTSSTFPMGSKSTIVLANTC 420
QY 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKPHSFHPATLISYLTAKPLVPPGTPVYN 480
DB 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKPHSFHPATLISYLTAKPLVPPGTPVYN 480

DB 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKPHSFHPATLISYLTAKPLVPPGTPVYN 480

QY 481 ALSKQPMLENTMRACVGLAPENNMLEYK 510

DB 481 ALSKQPMLENTMRACVGLAPENNMLEYK 510

RESULT 9

ADQ14503

ID ADQ14503 standard; protein; 510 AA.

AC ADQ14503;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #4.

KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme; mutant; mutain.

OS Glycine max.

XX Synthetic.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.

PR 11-MAR-2002; 2002US-00025003.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

PT WPI, 2004-533135/51.

PS N-PSDB; ADQ14502.

XX Example 8; SEQ ID NO 14; 48bp; English.

XX
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate synthase having a decreased capacity for the synthesis of myo-inositol 1-phosphate. The invention also relates to a chimeric gene operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant comprising the chimeric gene, a method of making the soybean plant, a seed of the soybean plant, a soy protein product derived from the processing of soybean seeds, a method of making or producing a soy protein product and a method of using a soybean plant homozygous for at least one gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate. The nucleic acid fragment and methods are useful for producing plants with decreased raffinose, stachyose and phytic acid content and increased sucrose and inorganic phosphate content, leading to valuable and useful soybean products. This sequence represents a mutant soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

CC Sequence 510 AA;

Query Match	98.8%	Score 2600;	DB 8;	Length 510;
Best Local Similarity	98.6%	Pred. No. 5.9e-210;		
Matches 503; Conservative	4;	Mismatches 3;	Indels 0;	Gaps 0;

Qy	1	MFLENKRECPNKKYETEISQSVYVNETTELVHENRNTQYQIKPKSVKKEPFTINHVP	60
Db	1	MFLENKRECVSBNKKYETEISQSVYNTEYELVHENRNGYQVIVAPKSVNVOFTINHVP	60
Qy	61	KLGVMVLVGMGNGNSLTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE	120
Db	61	KLGVMVLVGMGNGNSLTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE	120
Qy	121	IYAPFSSLLPMYNPDDIVFGWMDISMMNLADMARAKYFDDILQOKLAPVNESMLPLPGI	180
Db	121	IYAPFSSLLPMYNPDDIVFGWMDISMMNLADMARAKYFDDILQOKLAPVNESWPLPGI	180
Qy	181	YDADFPTAANOEBERANNVIGTKQEOVOOQIKDKAKFEATKVDKVVVLMTANTERSYNLV	240
Db	181	YDPDFPTAANOEBERANNVIGTKQEOVOOQIKDKAKFEATKVDKVVVLMTANTERSYNLV	240
Qy	241	VGANDPTMENLLAVDSENEAISPSTLYALACMEVNPFLNPSQNTPLPGLIDILAIRNT	300
Db	241	VGANDPTMENLLAVDSENEAISPSTLYALACMEVNPFLNPSQNTPLPGLIDILAIRNT	300
Qy	301	LIGGDDPFKSGQTKKRSVLVDPLVGAGIKRPTSIVSYNHLGNDGNMLSAPQTFERSKEISKS	360
Db	301	LIGGDDPFKSGQTKKRSVLVDPLVGAGIKRPTSIVSYNHLGNDGNMLSAPQTFERSKEISKS	360
Qy	361	NYVDMDMNSAALLIYEEGEBPDHVVVLYKYVPYIGDSKRAMDEYTSIIPFGSGSTVLANTC	420
Db	361	NYVDMDMNSAALLIYEEGEBPDHVVVLYKYVPYIGDSKRAMDEYTSIIPFGSGNTVLANTC	420
Qy	421	EDSLLAAPRIILDVYLLAELSTRILEFKAENEGFHSFHPAATILSVLTAPLVPPEGTVVN	480
Db	421	EDSLLAAPRIILDVYLLAELSTRIQFKAENEGFHSFHPAATILSVLTAPLVPPEGTVVN	480
Qy	481	ALSKORAMLENIRACVGLAPERNMTLEXY 510	
Db	481	ALSKORAMLENIRACVGLAPERNMTLEXY 510	

RESULT 10	
ADQ14505	ID
ADQ14505	ADQ14505 standard; protein; 510 AA.
XX	
AC	
ADQ14505;	
XX	
DT	23-SEP-2004 (first entry)
DE	Wild type soybean myo-inositol 1-phosphate synthase polypeptide #2.
XX	
KM	Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KM	raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme
XX	
OS	Glycine max.
XX	
PN	US2004128713-A1.
XX	
PD	01-JUL-2004.
XX	
PF	21-NOV-2003; 2003US-00718952.
XX	
PR	08-APR-1997; 97US-00835751.
PR	07-APR-1998; 98WO-US006822.
PR	26-APR-1999; 99US-00299315.
PR	11-MAR-2002; 2002US-00025003.
XX	
PA	(HITZ/) HITZ W D.
PA	(SEBA/) SEBASTIAN S A.
PA	(GRAC/) GRACE D J.
PA	(STRE/) STREIT L G.
XX	
HI	Hiltz WD, Sebastian SA, Grace DJ, Streit LG;

XX WP1; 2004-533135/51.
 DR N-PSDB; ADQ14504.
 DR
 XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytyc acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX
 XX Claim 3; SEQ ID NO 16; 48pp; English.
 PS

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a wild type soybean myo-
CC inositol 1-phosphate synthase polypeptide of the invention.
XX
90 Sequence 510 AA;

SQ Sequence 510 AA;

Query Match	98.8%	Score 2600;	DB 8;	Length 510;
Best Local Similarity	98.6%	Pred. No. 5.9e-210;		
Matches 503; Conservative	4;	Mismatches 3;	Indels 0;	Gaps 0

Qy	I	MFIENFKVBCPNVKTKTEIEIOSVYVYETTELVHENRGTQYIMKPKSVKYEKPTNIHP	60
Db	1	MFIENFKVSPNKKYTEIEIOSVYVYETTELVHENRGTQYIMKPKSVYQFKTNHP	60
Qy	61	KLGWMLVWGCGNNGSLTGGVIANREGISMATKOKIOQANFGSLDQASAIRGSGFCEE	120
Db	61	KLGWMLVWGCGNNGSLTGGVIANREGISMATKOKIOQANFGSLDQASAIRGSGFCEE	120
Qy	121	IYAPFSLPLPNVPPDIIVFGWDISNNMLADAMARAKFIDIDOKOLPYMESMLPLPGI	180
Db	121	IYAPFSLPLPNVPPDIIVFGWDISNNMLADAMARAKFIDIDOKOLRPYMESVPLPGI	180
Qy	181	YDPDFIAANOEBERANVIKGTQOEVOQIIXDKIAFKKATVDKVVYLMANTERYSNLV	240
Db	181	YDPDFIAANOEBERANVIKGTQOEVOQIIXDKIAFKKATVDKVVYLMANTERYSNLV	240
Qy	241	VGLNDPMENMLAAVDNBEATISPSLTLYAIACMEVPPINSPONTEVPGIDILATARN	300
Db	241	VGLNDPMENMLAAVDNBEATISPSLTLYAIACMEVPPINSPONTEVPGIDILATARN	300
Qy	301	LIGGDPFKSGQTKMSVLVDPLVGATIKPTSIIVSNHGNNDGNMLAPOTFRSKIEIS	360
Db	301	LIGGDPFKSGQTKMSVLVDPLVGATIKPTSIIVSNHGNNDGNMLAPOTFRSKIEIS	360
Qy	361	NVVDWMVNSNALLYEBGEHPDHVVIKTVPYVGSKRAMEYTSIEIMGGKSTIVLHNTC	420
Db	361	NVVDWMVNSNALLYEBGEHPDHVVIKTVPYVGSKRAMEYTSIEIMGGKSTIVLHNTC	420
Qy	421	EDSLLAAPITLIDLVLIAELSTRIEKKAENEGKFSFHPVATITLSYLTAKPLVPGPVYN	480
Db	421	EDSLLAAPITLIDLVLIAELSTRIOKKAENEGKFSFHPVATITLSYLTAKPLVPGPVYN	480
Qy	481	ALSKORAMENIMRACVGLAPENNMILTEYK	510
Db	481	ALSKORAMENIMRACVGLAPENNMILTEYK	510

Query	Match	Score	DB	Length
Beet Local Similarity	98.6%;	2600;	8;	510;
ADSB82004	standard;	510 AA.		
ADSB82004;				
ADSB82004;				
18-NOV-2004	(first entry)			
Soybean myo-inositol 1-phosphate synthase wild-type 4.				
Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid; raffinose; stachyose; sucrose; inorganic phosphate; flatulence.				
Glycine max; line 29018JPF03.				
US2003074685-A1.				
17-APR-2003.				
11-MAR-2002; 2002US-00025003.				
08-APR-1997; 97US-00835751.				
07-APR-1998; 98WO-US006822.				
(HITZ/) HITZ W D.				
(SEBA/) SEBASTIAN S A.				
Hitz WD, Sebastian SA;				
WPI: 2004-639857/62.				
N-P5DB; ADSB82003.				
Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.				
Example 8; SEQ ID NO 14; 34pp; English.				
The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homologous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a wild-type myo-inositol 1-phosphate synthase.				
Sequence 510 AA;				

	Matches	503; Conservative	4; Mismatches	3; Indels	0; Gaps	0; PT
Qy		1	MFENFKYECBNVXYTTEIQTISVYNYETTEL	VHNRNGTYQMIYKPKSVKYEFTKTNHVP	60	
Db		1	MFENFKYESNNVXYTTEIQTISVYNYETTEL	VHNRNGTYQMIYKPKSVNYQFTKTNHVP	60	
Qy		61	KLGVWLVMGNGNSSTLTGVIYANREGISWATKDKIQOANTYFGSLTQASAIRVSGFQEE	120		
Db		61	KLGVWLVMGNGNSSTLTGVIYANREGISWATKDKIQOANYFGSLTQASAIRVSGFQEE	120		
Qy		121	IYAPFKSLPLPVNPPDDIVFGGMDISNNMLADAMARAKVFDIDLKOLRPYMESMLPLPGI	180		
Db		121	IYAPFKSLPLPVNPPDDIVFGGMDISNNMLADAMARAKVFDIDLKOLRPYMESWVPLPGI	180		
Qy		181	YDPPFIANQBERANNVYKGTQKQEOVOQIIDIKAFKCATKVDKVVVLTMTANTERYSNLV	240		
Db		181	YDPPFIANQBERANNVYKGTQKQEOVOQIIDIKAFKCATKVDKVVVLTMTANTERYSNLV	240		
Qy		241	VGNDTMENLLAAVDNRNAEISPSSTLYAIACWMENVPPINGSPONTFVPGILIDIALANT	300		
Db		241	VGNDTMENLLAAVDNRNAEISPSSTLYAIACWMENVPPINGSPONTFVPGILIDIALANT	300		
Qy		301	LIGGDDFFSGGTTKKKSVLYVDFVAGTKPISIVSYNHLGNNDGNMLSAPOFRSKEAISKS	360		
Db		301	LIGGDDFFSGGTTKKKSVLYVDFVAGTKPISIVSYNHLGNNDGNMLSAPOFRSKEAISKS	360		
Qy		361	NVVDMDVNSNAILLYEPGEHPDHVVYIKVVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC	420		
Db		361	NVVDMDVNSNAILLYEPGEHPDHVVYIKVVPYVGSKRAMDEYTSIEIFMGKNTIVLHNTC	420		
Qy		421	EDSLIAPRIIDLVLALBELSTRIEFKAKENEGKFFSHFPAVITISLYTKAPLVPPGTVPVN	480		
Db		421	EDSLIAPRIIDLVLALBELSTRIOFKAKENEGKFFSHFPAVITISLYTKAPLVPPGTVPVN	480		
Qy		481	ALSKORAMLENIMRACVGLAENNMLLEYK	510		
Db		481	ALSKORAMLENIMRACVGLAENNMLLEYK	510		
RESULT 12						
ADSB2006	ID	ADSB2006	standard; protein; 510 AA.			
XX	AC	ADSB2006;				
XX	DT	18-NOV-2004	(first entry)			
DE	DE	Soybean myo-inositol 1-phosphate synthase wild-type 2.				
XX	KM	Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;				
KW	XX	raffinose; stachyose; sucrose; inorganic phosphate; flatulence.				
OS	XX	Glycine max; cultivar Wye.				
PN	XX	US2003074685-A1.				
PD	XX	17-APR-2003.				
XX	PF	11-MAR-2002; 2002US-00025003.				
PR	XX	08-APR-1997; 97US-00835751.				
PR	XX	07-APR-1998; 98WO-US006822.				
PA	XX	(HITZ/) HITZ W D.				
PA	XX	(SEBA/) SEBASTIAN S A.				
PI	XX	Hitz WD, Sebastian SA;				
DR	XX	WPI; 2004-639957/62.				
DR	XX	N-PSDB; ADSB2005.				
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose,					

PT phytic acid and inorganic phosphate content of soybean seeds.
XX
PS Claim 3; SEQ ID NO 16; 34pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis of myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.
SQ Sequence 510 AA:

Query Match 98.8%; Score 2600; DB 8; Length 510;
Best Local Similarity 98.6%; Pred. No. 5.9e-210;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MFIENPKVCEPNVKTETETIGSVNYETTELVEHNRNGTYOIVPKSVKEFKTNHVP 60
DB 1 MFIENPKVCEPNVKTETETIGSVNYETTELVEHNRNGTYOIVPKSVKEFKTNHVP 60
QY 1 MFIENPKVCEPNVKTETETIGSVNYETTELVEHNRNGTYOIVPKSVKEFKTNHVP 60
DB 1 MFIENPKVCEPNVKTETETIGSVNYETTELVEHNRNGTYOIVPKSVKEFKTNHVP 60
QY 61 KLGWLVGNGNGSTLTGVIANREGISWATKDKIQOANNYGSITQASAIRVGSFOGEE 120
DB 61 KLGWLVGNGNGSTLTGVIANREGISWATKDKIQOANNYGSITQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPWNPDPDIVFGWMDISNMNLADAMAKVYEDIDLOKQLRPYMSMLPLPGI 180
DB 121 IYAPFKSLIPWNPDPDIVFGWMDISNMNLADAMAKVYEDIDLOKQLRPYMSMLPLPGI 180
QY 181 YDPDFIAANQEEERANNVTKGTQEOVOQIIKOIKAPKEATKYDKVYVMTANTERYSNLY 240
DB 181 YDPDFIAANQEEERANNVTKGTQEOVOQIIKOIKAPKEATKYDKVYVMTANTERYSNLY 240
QY 181 YDPDFIAANQEEERANNVTKGTQEOVOQIIKOIKAPKEATKYDKVYVMTANTERYSNLY 240
DB 181 YDPDFIAANQEEERANNVTKGTQEOVOQIIKOIKAPKEATKYDKVYVMTANTERYSNLY 240
QY 241 VGLNMTMNLAAVDRNEAISPSTLYAACWENVPFINSPOPTVPGGLIDLAIAANT 300
DB 241 VGLNMTMNLAAVDRNEAISPSTLYAACWENVPFINSPOPTVPGGLIDLAIAANT 300
QY 301 LIIGDDPKSGQTKMKSVLVDFVGAGIKPSTISVYNNHGNNGMMLASAPQFRSKESKS 360
DB 301 LIIGDDPKSGQTKMKSVLVDFVGAGIKPSTISVYNNHGNNGMMLASAPQFRSKESKS 360
QY 361 NVVDDMVNSNALIYRGEHPDHVVVVKYVPYGDSCRADDEYTSSEIFMGKSTIYLAHTC 420
DB 361 NVVDDMVNSNALIYRGEHPDHVVVVKYVPYGDSCRADDEYTSSEIFMGKSTIYLAHTC 420
QY 421 EDSLLAAPTILDLVLAELSTRIFPKAENEGKFSHPATLSTLYTARAPLVPGTPVNV 480
DB 421 EDSLLAAPTILDLVLAELSTRIFPKAENEGKFSHPATLSTLYTARAPLVPGTPVNV 480
SQ Query Match 98.8%; Score 2600; DB 8; Length 510;
Best Local Similarity 98.6%; Pred. No. 5.9e-210;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 481 ALSKQRAMLENIMRACVGLADENNMILEYK 510
DB 481 ALSKQRAMLENIMRACVGLADENNMILEYK 510
RESULT 13
ADQ14501
ID ADQ14501 standard; protein, 510 AA.
XX
AC ADQ14501;
XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #3.
XX
KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KW mutant; mutant.
XX
OS Glycine max.
OS Synthetic.
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US0006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
DR WPI, 2004-533135/51.
XX
DR N-PSDB; ADQ14500.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Claim 9; SEQ ID NO 12; 48pp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.
SQ Sequence 510 AA:

Query Match 98.5%; Score 2593; DB 8; Length 510;
Best Local Similarity 98.4%; Pred. No. 2.3e-209;

Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MEIENFKVCEPVKKTETETISQSVNYVETTELVEHNRNGTYQVIKPKSVKTEPKTNINHP 60
DB 1 MEIENFKVESPVKKTETETISQSVNYVETTELVEHNRNGTYQVIKPKSVNYOFTKTNHVP 60
QY 61 KLGWLVGMGNGNSTLTGCVIANREGISMAWKDKIQOANYGSLTQASAIRVGSFGQEE 120
DB 61 KLGWLVGMGNGNSTLTGCVIANREDISMAWKDKIQOANYGSLTQASAIRVGSFGQEE 120
QY 121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOBERANNAVIGTKQEOVOQIIOIKAFKFKATKVDKVVLTMTANTERSNLY 240
DB 181 YDPDFIAANOBERANNAVIGTKQEOVOQIIOIKAFKFKATKVDKVVLTMTANTERSNLY 240
QY 241 VGLNDTMENTLAAVDRNEAEISPTLYAIACVMENVPFINGSPONTFVPGILDIAIANT 300
DB 241 VGLNDTMENTLAAVDRNEAEISPTLYAIACVMENVPFINGSPONTFVPGILDIAIANT 300
QY 301 LIGDDPFKSGQTKMSVLVDPLVGAGIKPTSTIVSYNHLGNNDGMNLAPQTRSKSISKS 360
DB 301 LIGDDPFKSGQTKMSVLVDPLVGAGIKPTSTIVSYNHLGNNDGMNLAPQTRSKSISKS 360
QY 361 NVYDDMVNSNALIYERGEHPDHVVVIKVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVYDDMVNSNALIYERGEHPDHVVVIKVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPTILDLVLAELSTRILEPKAENEGKPHSPATILSYLTKAPLVPPTPVVN 480
DB 421 EDSLLAAPTILDLVLAELSTRILEPKAENEGKPHSPATILSYLTKAPLVPPTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 14
ADS82002
ID ADS82002 standard; protein; 510 AA.
XX
AC ADS82002;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase mutant #2.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
KM phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
KM flatulence; mutant.
XX
OS Glycine max; line 29010CP01.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 87
FT /note= "wild-type Gly substituted by Asp"
XX
PN US2003074685-A1.
PD 17-APR-2003.
PE 11-MAR-2002; 2002US-00025003.
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
PI Hitz WD, Sebastian SA;

XX MPI: 2004-639957/62.
DR N-PSDB; ADS82001.
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
PS Claim 9; SEQ ID NO 12; 34pp; English.
XX
The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the gene synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a mutant myo-inositol 1-phosphate synthase.

SO Sequence 510 AA;
Query Match 98.5%; Score 2593; DB 8; Length 510;
Best Local Similarity 98.4%; Pred. No. 2.3e-209;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MEIENFKVCEPVKKTETETISQSVNYVETTELVEHNRNGTYQVIKPKSVKTEPKTNINHP 60
DB 1 MEIENFKVESPVKKTETETISQSVNYVETTELVEHNRNGTYQVIKPKSVNYOFTKTNHVP 60
QY 61 KLGWLVGMGNGNSTLTGCVIANREGISMAWKDKIQOANYGSLTQASAIRVGSFGQEE 120
DB 61 KLGWLVGMGNGNSTLTGCVIANREDISMAWKDKIQOANYGSLTQASAIRVGSFGQEE 120
QY 121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOBERANNAVIGTKQEOVOQIIOIKAFKFKATKVDKVVLTMTANTERSNLY 240
DB 181 YDPDFIAANOBERANNAVIGTKQEOVOQIIOIKAFKFKATKVDKVVLTMTANTERSNLY 240
QY 241 VGLNDTMENTLAAVDRNEAEISPTLYAIACVMENVPFINGSPONTFVPGILDIAIANT 300
DB 241 VGLNDTMENTLAAVDRNEAEISPTLYAIACVMENVPFINGSPONTFVPGILDIAIANT 300
QY 301 LIGDDPFKSGQTKMSVLVDPLVGAGIKPTSTIVSYNHLGNNDGMNLAPQTRSKSISKS 360
DB 301 LIGDDPFKSGQTKMSVLVDPLVGAGIKPTSTIVSYNHLGNNDGMNLAPQTRSKSISKS 360
QY 361 NVYDDMVNSNALIYERGEHPDHVVVIKVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVYDDMVNSNALIYERGEHPDHVVVIKVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420

```

Db 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGCKNTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKAPLVPFGTPPVN 480
Db 421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPFGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 15

AA24477
ID AA24477 standard; protein; 536 AA.

XX AC AA24477;

DT 24-SEP-1999 (first entry)

XX Nicotiana paniculata INPS protein.

XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;

KW water stress; resistance.

XX Nicotiana paniculata.

OS Nicotiana paniculata.

XX Key Location/Qualifiers

FT Misc-difference 511 /label= unknown

FT /note= "encoded by the stop codon TGA"

PN JP1187879-A.

XX 13-JUL-1999.

XX 26-DEC-1997; 97JP-00359773.

XX 26-DEC-1997; 97JP-00359773.

XX (NISB) JAPAN TOBACCO INC.

XX WPI, 1999-451546/38.

DR N-PSDB; AAX90402.

PT New INPS gene derived from Nicotiana genus plant - useful for conferring resistance to water stress to plants.

XX Claim 2; Page 6-8; Bpp; Japanese.

CC The present sequence is the Nicotiana paniculata inositol monophosphate synthase (INPS), designated NpINPS1. INPS can be used to confer water stress resistance to a plant

XX Sequence 536 AA;

Query Match 93.9%; Score 2472; DB 2; Length 536;

Best Local Similarity 92.2%; Pred. No. 3.8e-199;

Matches 470; Conservative 24; Mismatches 16; Indels 0; Gaps 0;

QY 1 MEIENFKVECPNVKTYETIEIQSVINYEETELVHENRNGTYOMIVKPKSVKYEFTNIHVP 60
Db 1 MEIENFKVESPMVKYTESRHSVYDYQTELHDEKNGTYQWTVKPKTVKYEFTKTDVHVP 60
QY 61 KLGVMVLGVGGNGSTLNGVIANREGISWATKDKIQOANVFGSLTQASATRVGSFQGE 120
Db 61 KLGVMVLGVGGNGSTLNGVIANREGISWATKDKIQOANVFGSLTQASTIRVGSFNGEE 120
QY 121 IYAPFKSLPLMVPNPDIVFGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFKSLPLMVPNPDIVFGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMPLPGI 180
QY 181 YDPDFIAANQERRANNVIGTKYQBOVQOIIKDIKAFKEATKVDKRVVLTANTERYSNLV 240
Db 181 YDPDFIAANQERRANNVIGTKYQBOVQOIIKDIKAFKEATKVDKRVVLTANTERYSNLV 240

Db 181 YDPDFIAANQERRANNVIGTKYQBOVQOIIKDIKAFKEATKVDKRVVLTANTERYSNLV 240
QY 241 VGLNDTMENLAAVDRNEAEISPSLTVAIACWMENVPFINGSPONTFVPGIIDLAIANT 300
Db 241 VGLNDTMENLAAVDRNEAEISPSLTVAIACLENVPFINGSPONTFVPGIIDLAIKNT 300
QY 301 LIGDDFKSGGOTKXKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLAPOTFRSKEISKS 360
Db 301 LIGDDFKSGGOTKXKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLAPOTFRSKEISKS 360
QY 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGCKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGCKNTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKAPLVPFGTPPVN 480
Db 421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPFGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

Search completed: June 7, 2005, 16:43:36
Job time : 99.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:16:29 ; Search time 24.5 Seconds
(without alignments)
2002.880 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MFIEHFKECPNVKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 79:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	2381	90.5	512	2 T12438	inositol-3-phospha
2	2378	90.3	510	2 D84610	probable myo-inos
3	2369	90.0	510	2 T50021	inositol-3-phospha
4	2363	89.8	509	2 T08436	inositol-3-phospha
5	2352.5	89.4	511	2 T05017	inositol-3-phospha
6	2349	89.2	510	2 S60302	inositol-3-phospha
7	2344	89.1	510	2 T01647	inositol-3-phospha
8	2331.5	88.6	511	2 T10964	inositol-3-phospha
9	2328	88.4	510	2 T04399	inositol-3-phospha
10	2314.5	87.9	507	2 S52648	inositol-3-phospha
11	1386.5	52.8	525	2 T18569	inositol-3-phospha
12	1382.5	52.5	555	2 A30902	inositol-3-phospha
13	1381.5	52.5	520	2 S45452	inositol-3-phospha
14	1308.5	49.7	430	2 T46317	inositol-3-phospha
15	296	11.2	388	2 T36191	probable myo-inosi
16	285.5	10.8	392	2 A69474	myo-inositol-1-pho
17	242	9.2	417	2 T34930	probable secreted
18	214	8.1	382	2 H72255	myo-inositol-1-pho
19	212	8.1	425	2 E71039	hypothetical prote
20	198	7.5	386	2 B75175	myo-inositol-1-pho
21	195	7.4	364	2 G70451	conserved hypothet
22	163	6.2	360	2 B90239	conserved hypothet
23	158	6.0	392	2 F72632	probable myo-inosi
24	145.5	5.5	368	2 A65014	conserved hypothet
25	142	5.4	369	2 S72835	hypothetical prote
26	133	5.1	360	2 T36586	hypothetical prote
27	128.5	4.9	367	2 F70912	hypothetical prote
28	125.5	4.8	367	2 G95871	conserved hypothet
29	119.5	4.5	873	2 T16282	hypothetical prote

30	118	4.5	739	2 B86816	hypothetical prote
31	116.5	4.4	244	2 B90176	conserved hypothet
32	116	4.4	535	2 A80983	periplasmic dipepc
33	115.5	4.4	731	2 D95235	penicillin-binding
34	115.5	4.4	1993	2 A71450	probable peptidogl
35	115	4.4	2490	1 A54971	protein-tyrosine-p
36	114.5	4.4	867	2 E86815	ClpB protein (limo
37	114.5	4.4	900	2 S46101	hypothetical prote
38	113	4.3	351	2 D82930	Srp family of GTP-
39	113	4.3	793	2 T51701	phosphoribosylform
40	113	4.3	802	2 C90326	hypothetical prote
41	113	4.3	4936	2 A42515	hypothetical prote
42	112.5	4.3	430	2 T04668	phosphoserine tran
43	112.5	4.3	867	2 H90524	preprotein translo
44	111.5	4.2	731	2 B98099	peptidoglycan glyc
45	111.5	4.2	776	2 S59790	hypothetical prote

ALIGNMENTS

RESULT 1

T12438
inositol-3-phosphate synthase (EC 5.5.1.4) - common ice plant
C.Species: Mesembryanthemum crystallinum (common ice plant)
C.Date: 23-Jul-1999 #sequence #revision 23-Jul-1999 #text_change 09-Jul-2004
C.Accession: T12438
R.Ishitani, M.; Majumder, A.L.; Bornhauser, A.; Michalowski, C.B.; Jensen, R.G.; Bohnert, Plant J. 9, 537-548, 1996
A.Title: Coordinate transcriptional induction of myo-inositol metabolism during enviro
A.Reference number: Z17518; MUID:96208959; PMID:8624516
A.Accession: T12438
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-512 <ISH>
A.Cross-references: UNIPROT:Q40271; EMBL:U32511; NID:g975887; PION:AA03687.1; PID:g97588
C.Superfamily: myo-inositol-1-phosphate synthase
C.Keywords: intramolecular lyase; isomerase

Query Match	90.5%; Score 2381; DB 2; Length 512;
Best Local Similarity	88.7%; Pred. No. 4.6e-153;
Matches	454; Conservative 31; Mismatches 25; Indels 2; Gaps 1;
QY	1 MFIEHFKECPNVKXTETETIETELVHKN--GTYQIVKRSVKYERKTH 58
DB	1 MFIEHFKECPNVKXTETETIETELVHKN--GTYQIVKRSVKYERKTH 60
QY	59 VKLGVMLVGMGNGNSTLTGVIANRGISWATKDIQANYPGSLTQASIRVGFQ 118
DB	61 VKLGVMLVGMGNGNSTLTGVIANRGISWATKDIQANYPGSLTQASIRVGFQ 120
QY	119 BEIYAFKSLPMPVDDIVFGMDISNMNLADAMARAKVPDIDLOKLRPMESMLPLP 178
DB	121 BEIYAFKSLPMPVDDIVFGMDISNMNLADAMARAKVPDIDLOKLRPMESMLPLP 180
QY	179 GIYDPDFIAANDPEFANNVYKTKOEYQOIKIKAFKREATKDVRYVLTANTERYSN 238
DB	181 GIYDPDFIAANDPEFANNVYKTKOEYQOIKIKAFKREATKDVRYVLTANTERYSN 240
QY	239 LVVGIINDMENLAVDNEAEISPTLYATAQWENPFINGSQNTFFVPELIDLAIR 298
DB	241 LVVGIINDMENLAVDNEAEISPTLYATAQWENPFINGSQNTFFVPELIDLAIR 300
QY	299 NTLIGDDPFKSGQTQKKSVLVDPLVAGIKPSTISVYHNLGNDGMSAPQTFPSKES 358
DB	301 NSLIGDDPFKSGQTQKKSVLVDPLVAGIKPSTISVYHNLGNDGMSAPQTFPSKES 360
QY	359 KSNVVDMAVNSNAILYEPGEHPDHVVVYKYPYVDSFRANDYTSSEIFMGSKTIVLHN 418
DB	361 KSNVVDMAVNSNAILYEPGEHPDHVVVYKYPYVDSFRANDYTSSEIFMGSKTIVLHN 420
QY	419 TCEDSLAAPITLIDVLAELSTRIEFKAEKNGKSHSPAVATLISYTKAPLVPRGTPV 478

Db 421 TCEDSLAAPILIDVLVLAELSTRIQLKAEEDKFSFHPVATILSYLTKAPLVPPGTPV 480
Qy 479 VNALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 VNALSKORAMLENIMRACVGLAPENNMILEYK 512

RESULT 2
D84610
probable myo-inositol 1-phosphate synthase [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004
C/Accession: D84610
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nieman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-766, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: D84610
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-510 <STO>
A/Cross-references: UNIPROT:Q38662; GB:AE002093; NID:g4567202; PIDN:AAD33618.1; GSPDB:GN
A/Map position: 2
C/Superfamily: Myo-inositol-1-phosphate synthase

Query Match 90.3%; Score 2378; DB 2; Length 510;
Best Local Similarity 87.8%; Pred. No. 7.3e-153;
Matches 448; Conservative 40; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MFIENFKVCEPNVKKYTEIETETIOSVYNYETTELVHENRNGTYQWIVPKSKVKEFKTNHVP 60
Db 1 MFISFKVESPNVKKTENISVVDYQTTTELVHENKNGAFQWTVKPKVKEFKTDTRVP 60

Qy 61 KLGVWLVMGNGNGSTLTGCVIANREGISMAWKIKQANVFGLSTQASIRVGSFOGEE 120
Db 61 KLGVWLVMGNGNGSTLTGCVIANREGISMAWKIKQANVFGLSTQASIRVGSFOGEE 120

Qy 121 IYAPFKSLIPVNPDPDIYFGGMDISNNMLADAMARAKVFDIDLOKQAPRYMESMLPLPGI 180
Db 121 IYAPFKSLIPVNPDPDIYFGGMDISNNMLADAMARAKVFDIDLOKQAPRYMENMLPLPGI 180

Qy 181 YAPDFIAANQSSRAHNVIKGTKEOVQOIIIDIKAFKEATYDKVYVLTANTERYSNLV 240
Db 181 YAPDFIAANQSSRAHNVIKGTKEOVQOIIIDIKAFKEATYDKVYVLTANTERYSNLV 240

Qy 241 VGLNDTMENTLAAVDRNEAETSPSTLYAACVMEVNPFGNSPONTFVPGILDLAIART 300
Db 241 VGLNDTMENTLAAVDRNEAETSPSTLYAACVMEVNPFGNSPONTFVPGILDLAIART 300

Qy 301 LIIGDDFKSGQTKMSVLDVFLVAGAIKPTISIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 301 LIIGDDFKSGQTKMSVLDVFLVAGAIKPTISIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360

Qy 361 NVDDMVNSNALIYEPGEHPDHVVVYKYVPYVDSKRAMDEYTSSEIFMGKSTIYLANTC 420
Db 361 NVDDMVNSNALIYEPGEHPDHVVVYKYVPYVDSKRAMDEYTSSEIFMGKSTIYLANTC 420

Qy 421 EDSLAAPIIIDVLVLAELSTRIEFKANEKGKFSFHPVATILSYLTKAPLVPPGTPV 480
Db 421 EDSLAAPIIIDVLVLAELSTRIEFKANEKGKFSFHPVATILSYLTKAPLVPPGTPV 480

Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 3
T50021
inositol-3-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis thaliana

N/Alternate names: protein T31P16.160
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C/Accession: T50021
R/Beyan, M.; Zimmermann, W.; Grennaisen, A.; Wambut, R.; Kalicki, J.; Wohldmann, P.; Smi
submitted to the Protein Sequence Database, May 2000
A/Reference number: 225027
A/Accession: T50021
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-510 <BEV>
A/Cross-references: UNIPROT:Q9LX12; EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.160
A/Experimental source: cultivar Columbia; BAC clone T31P16
C/Genetics:
A/Map position: 5
A/Intons: 63/2; 86/2; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 90.0%; Score 2369; DB 2; Length 510;
Best Local Similarity 87.5%; Pred. No. 3e-152;
Matches 446; Conservative 42; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MFIENFKVCEPNVKKYTEIETETIOSVYNYETTELVHENRNGTYQWIVPKSKVKEFKTNHVP 60
Db 1 MFISFKVESPNVKKTENISVVDYQTTTELVHENKNGAFQWTVKPKVKEFKTDTRVP 60

Qy 61 KLGVWLVMGNGNGSTLTGCVIANREGISMAWKIKQANVFGLSTQASIRVGSFOGEE 120
Db 61 KLGVWLVMGNGNGSTLTGCVIANREGISMAWKIKQANVFGLSTQASIRVGSFOGEE 120

Qy 121 IYAPFKSLIPVNPDPDIYFGGMDISNNMLADAMARAKVFDIDLOKQAPRYMESMLPLPGI 180
Db 121 IYAPFKSLIPVNPDPDIYFGGMDISNNMLADAMARAKVFDIDLOKQAPRYMENMLPLPGI 180

Qy 181 YAPDFIAANQSSRAHNVIKGTKEOVQOIIIDIKAFKEATYDKVYVLTANTERYSNLV 240
Db 181 YAPDFIAANQSSRAHNVIKGTKEOVQOIIIDIKAFKEATYDKVYVLTANTERYSNLV 240

Qy 241 VGLNDTMENTLAAVDRNEAETSPSTLYAACVMEVNPFGNSPONTFVPGILDLAIART 300
Db 241 VGLNDTMENTLAAVDRNEAETSPSTLYAACVMEVNPFGNSPONTFVPGILDLAIART 300

Qy 301 LIIGDDFKSGQTKMSVLDVFLVAGAIKPTISIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 301 LIIGDDFKSGQTKMSVLDVFLVAGAIKPTISIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360

Qy 361 NVDDMVNSNALIYEPGEHPDHVVVYKYVPYVDSKRAMDEYTSSEIFMGKSTIYLANTC 420
Db 361 NVDDMVNSNALIYEPGEHPDHVVVYKYVPYVDSKRAMDEYTSSEIFMGKSTIYLANTC 420

Qy 421 EDSLAAPIIIDVLVLAELSTRIEFKANEKGKFSFHPVATILSYLTKAPLVPPGTPV 480
Db 421 EDSLAAPIIIDVLVLAELSTRIEFKANEKGKFSFHPVATILSYLTKAPLVPPGTPV 480

Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 4
T08436
inositol-3-phosphate synthase (EC 5.5.1.4) [similarity] - rape
C/Species: Brassica napus (rape)
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: T08436
R/Hussein, A.; Bourgeois, J.; Polvi, S.; Tsang, E.; Keller, W.A.; Georges, F.
submitted to the EMBL Data Library, August 1996
A/Reference number: Z16418
A/Accession: T08436
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA

A:Residues: 1-509 <HUS>
A:Cross-references: UNIPROT:Q96348; EMBL:U66307; NID:g1513227; PID:g1513228
C:Function:
A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 89.8%; Score 2363; DB 2; Length 509;
Best Local Similarity 88.0%; Pred. No. 7.5e-152;
Matches 448; Conservative 33; Mismatches 28; Indels 0; Gaps 0;

QY 2 FIENFKVCEPNVKTETETIOSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTNINHYV 61
DB 1 FIESFVSEPNVKTETNEISHVYDYETTEVVENHNGAYQWIVKPKVYKDTKIDRVK 60
QY 62 LGVNLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGEI 121
DB 61 LGVNLVGMGNGNSTLTGVIANKEGISWATKDKVQOANYFGSLTQASSIRVSGFNGEM 120
QY 122 VAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKOLRPYMESMLPLPGIY 181
DB 121 VAPFKSLPMVNPDDIVFGGWDISDNMLADAMGARAVDIDLOKOLRPYMENTVPLPGIY 180
QY 182 DDPFLAANQEBRANNVYKGTQKQOVQIITDKAFKATKVDKVVVLTANTERYSNLVY 241
DB 181 DDPFLAANQSRANNVYKGTKEQVDQIITKDMREFEKKNVDKVVVLTANTERYSNVY 240
QY 242 GLNDTMENTLAAVDNRNAEISPTLYAIACMENVPFINGSPQNTFVPGILDIAIRNLT 301
DB 241 GLNDTMENTLANSVDNRSEISPTLYAIACVLEGIFINGSPQNTFVPGILDIAIRNLT 300
QY 302 IGGDDPKSGQTKKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISK 361
DB 301 IGGDDPKSGQTKKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISK 360
QY 362 VVDDMNNSNALIYEPGEHPDHVVYIKVPPVGDGSKRAMDEYTSIEFMGKSTYLTANT 421
DB 361 VVDDMNVASNGILFEPGEHPDHVVYIKVPPVADSKRAMDEYTSIEFMGKSTYLTANT 420
QY 422 DSLAAPILIDVLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPGPSTVNA 481
DB 421 DSLAAPILIDVLVLAELSTRIOFKSEKGFSPHPVATKSYLTAKPLVPGPSTVNA 480
QY 482 LSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 LSKORAMLENIMRACVGLAPENNMILEYK 509

RESULT 5
T05017
inositol-3-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thalian
N:Alternate names: protein T19P19.190
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05017
R:Byvan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hoheisel, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15394
A:Accession: T05017
A:Molecule type: DNA
A:Residues: 1-511 <BRV>
A:Cross-references: UNIPROT:P42801; EMBL:A1022605
A:Experimental source: cultivar Columbia; BAC clone T19P19
C:Genetics:
A:Map position: 4
A:Introns: 64/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3
A>Note: T19P19.190
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 89.4%; Score 2352.5; DB 2; Length 511;
Best Local Similarity 87.7%; Pred. No. 3.9e-151;
Matches 448; Conservative 36; Mismatches 26; Indels 1; Gaps 1;

QY 1 MEIENFKVCEPNVKTETETIOSVYNYETTELVEHNR-NGTYQWIVKPKSVKYEFTNINHY 59
DB 1 MEIESFKVSEPNVKTETNEISHVYDYETTEVHEKTVNTQWIVKPKVYKDTKIDRV 60
QY 60 PKLGWNLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 119
DB 61 PKLGWNLVGMGNGNSTLTGVIANKEGISWATKDKVQOANYFGSLTQASSIRVSGFNGE 120
QY 120 EIVAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKOLRPYMESMLPLPG 179
DB 121 EIVAPFKSLPMVNPDDIVFGGWDISDNMLADAMARAVFDIDLOKOLRPYMENTVPLPG 180
QY 180 IYDPPFLAANQEBRANNVYKGTQKQOVQIITDKAFKATKVDKVVVLTANTERYSNLV 239
DB 181 IYDPPFLAANQSRANNVYKGTKEQVDHITKDMREFEKKNVDKVVVLTANTERYSNV 240
QY 240 VVGLNDTMENTLAAVDNRNAEISPTLYAIACMENVPFINGSPQNTFVPGILDIAIRN 299
DB 241 VVGLNDTMENTLANSVDNRDAEISPTLYAIACVLEGIFINGSPQNTFVPGILDIAIRN 300
QY 300 TLIIGDDPKSGQTKKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISK 359
DB 301 TLIIGDDPKSGQTKKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISK 360
QY 360 SNVDDMNNSNALIYEPGEHPDHVVYIKVPPVGDGSKRAMDEYTSIEFMGKSTYLTANT 419
DB 361 SNVDDMNVASNGILFEPGEHPDHVVYIKVPPVADSKRAMDEYTSIEFMGKSTYLTANT 420
QY 420 CEDSLAAPILIDVLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPGPSTPV 479
DB 421 CEDSLAAPILIDVLVLAELSTRIOFKSEKGFSPHPVATILSYLTAKPLVPGPSTPV 480
QY 480 NALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 NALSKORAMLENIMRACVGLAPENNMILEYK 511

RESULT 6
S60302
inositol-3-phosphate synthase (EC 5.5.1.4) - Spirodela polyrrhiza
C:Species: Spirodela polyrrhiza
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60302
R:Smart, C.C.; Fleming, A.J.
Plant J. 4, 279-293, 1993
A:Title: A plant gene with homology to D-myo-inositol-3-phosphate synthase is rapidly an
A:Reference number: S60302; MUID:94035182; PMID:8220483
A:Accession: S60302
A:Molecule type: mRNA
A:Residues: 1-510 <SMA>
A:Cross-references: UNIPROT:P42803; EMBL:Z11693; NID:G396209; PID:CAA7751.1; PID:G55864
C:Genetics:
A:Gene: turl
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 89.2%; Score 2349; DB 2; Length 510;
Best Local Similarity 87.3%; Pred. No. 6.6e-151;
Matches 445; Conservative 40; Mismatches 25; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPNVKTETETIOSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTNINHY 60
DB 1 MEIEKFRVSEPNVKTGDEIEISVYETTEVHEVNGSYQWVVRKSVQVOFKTIDRV 60
QY 61 KLGWNLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
DB 61 KLGWNLVGMGNGNSTLTGVIANREGISWATKEXQOANYFGSLTQASSIRVSGFNGE 120
QY 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLPMVNPDEIVFGGWDISDNMLADAMGARAVDIDLOKOLRPYMESMLPLPGI 180


```
Db 421 CEDSLAARPIIDLVLLAEISTRIQKSGEGKFSFHPVATILSYLTKAPLVPPTGPV 480
Qy 480 NALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 NALSKORAMLENIMRACVGLAPENNMILEYK 511

RESULT 9
T04399
Inositol-3-phosphate synthase (EC 5.5.1.4) - barley
C/Species: Hordeum vulgare (barley)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T04399
R/Larson, S.R.; Raboy, V.
Submitted to the EMBL Data Library, March 1998
A/Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A/Reference number: Z14366
A/Accession: T04399
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-510 <L&R>
A/Cross-references: UNIPROT:O65195; EMBL:AF056325; NID:g3152730; PIDN:AC17133.1; PID:g3
A/Experimental source: cv. Harrington
C/Genetics:
A:Gene: INO1
A:Map position: 4
C/Function:
A/Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inc
A/Pathway: inositol biosynthesis
A/Note: first step
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 88.4%; Score 2328; DB 2; Length 510;
Best Local Similarity 87.6%; Pred. No. 1.7e-149;
Matches 447; Conservative 26; Mismatches 37; Indels 0; Gaps 0;

Qy 1 MEIENPKVCEPVKKTETETIQSVNYETTELHNRNGTYOMIVKPKSVKYEKTNHVP 60
Db 1 MEISFVRESPPVRRGAGIESEYRDTELHVESHDGASKVVRKSVNYHFKTNTPV 60
Qy 61 KLGVMLVGVGNGNGSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQEE 120
Db 61 KLGVMLVGVGNGNGSTLTGMAGVIANREGISWATKDKIQOANYFGSLTQASTIRGSGNEE 120
Qy 121 IYAPFKSLIPVNPDPDIVFGWDISNNLADAMARAFVDIDLQKQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPVNPDPDIVFGWDISNNLADAMATRAKVDIDLQKQLRPYMESMPLPGI 180
Qy 181 YPDPFLAANQERANNVIGTKOEVOQIIXDKAFKEATKYDKVYVLTANTERSNLY 240
Db 181 YPDPFLAANQERANNVIGTKOEVOQIIXDKAFKEATKYDKVYVLTANTERSNLY 240
Qy 241 YPDPFLAANQERANNVIGTKOEVOQIIXDKAFKEATKYDKVYVLTANTERSNLY 240
Db 241 YPDPFLAANQERANNVIGTKOEVOQIIXDKAFKEATKYDKVYVLTANTERSNLY 240
Qy 241 VGLNTMENTLAAVDRNEAELSPSTLYAIACVMEVNPFGNSPONTFVGLIDLAIARNT 300
Db 241 VGLNTMENTLAAVDRNEAELSPSTLYAIACVMEVNPFGNSPONTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMILSAPOTRSKSISKS 360
Db 301 LIGGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMILSAPOTRSKSISKS 360
Qy 361 NVVDDMVNSNALYERBEHPDHVVVIKYVPYVGDSTRADDEYTSRIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALYERBEHPDHVVVIKYVPYVGDSTRADDEYTSRIFMGKSTIVLHNTC 420
Qy 421 EDSLAAARPIIDLVLLAEISTRIEFAENEGKFSFHPVATILSYLTKAPLVPPTGPV 480
Db 421 EDSLAAARPIIDLVLLAEISTRIEFAENEGKFSFHPVATILSYLTKAPLVPPTGPV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALAKORAMLENIMRACVGLAPENNMILEYK 510
```

```
RESULT 10
S52648
Inositol-3-phosphate synthase (EC 5.5.1.4) - Citrus paradisi
C/Species: Citrus paradisi
C/Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C/Accession: S52648
R/Holland, D.
Submitted to the EMBL Data Library, April 1994
A/Reference number: S52648
A/Accession: S52648
A/Molecule type: DNA
A/Residues: 1-507 <HOL>
A/Cross-references: UNIPROT:P42802; GB:Z32632; NID:g602564; PIDN:CAA83565.1; PID:g602565
C/Genetics:
A:Gene: INO1
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase

Query Match 87.9%; Score 2314.5; DB 2; Length 507;
Best Local Similarity 88.0%; Pred. No. 1.4e-148;
Matches 449; Conservative 30; Mismatches 28; Indels 3; Gaps 2;

Qy 1 MEIENPKVCEPVKKTETETIQSVNYETTELHNRNGTYOMIVKPKSVKYEKTNHVP 60
Db 1 MEIENPKVCEPVKKTETETIHSHVYDETTELHNRNGTYOMIVKPKSVKYEKTDVHP 60
Qy 61 KLGVMLVGVGNGNGSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQEE 120
Db 61 KLGVMLVGVGNGNGSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGNEE 120
Qy 121 IYAPFKSLIPVNPDPDIVFGWDISNNLADAMARAFVDIDLQKQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPVNPDPDIVFGWDISNNLADAMARAFVDIDLQKQLRPYMESMPLPGI 180
Qy 181 YPDPFLAANQERANNVIGTKOEVOQIIXDKAFKEATKYDKVYVLTANTERSNLY 240
Db 181 YPDPFLAANQERANNVIGTKOEVOQIIXDKAFKEATKYDKVYVLTANTERSNLY 240
Qy 241 YPDPFLAANQERANNVIGTKOEVOQIIXDKAFKEATKYDKVYVLTANTERSNLY 240
Db 241 YPDPFLAANQERANNVIGTKOEVOQIIXDKAFKEATKYDKVYVLTANTERSNLY 240
Qy 241 VGLNTMENTLAAVDRNEAELSPSTLYAIACVMEVNPFGNSPONTFVGLIDLAIARNT 300
Db 241 VGLNTMENTLAAVDRNEAELSPSTLYAIACVMEVNPFGNSPONTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMILSAPOTRSKSISKS 360
Db 301 LIGGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMILSAPOTRSKSISKS 360
Qy 361 NVVDDMVNSNALYERBEHPDHVVVIKYVPYVGDSTRADDEYTSRIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSN--VFMGLVNTRPRIKYVPYVA--IERADEYTSRIFMGKSTIVLHNTC 417
Qy 421 EDSLAAARPIIDLVLLAEISTRIEFAENEGKFSFHPVATILSYLTKAPLVPPTGPV 480
Db 421 EDSLAAARPIIDLVLLAEISTRIQKSGEGKFSFHPVATILSYLTKAPLVPPTGPV 477
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 478 ALSKORAMLENIMRACVGLAPENNMILEYK 507

RESULT 11
T18569
Inositol-3-phosphate synthase (EC 5.5.1.4) - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18569; T20002
R/Ainscough, R.
Submitted to the EMBL Data Library, December 1998
A/Reference number: T18569
A/Accession: T18569
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
```


Best Local Similarity 27.7%; Pred. No. 1.8e-12;
Matches 110; Conservative 62; Mismatches 157; Indels 68; Gaps 11;

```
OY      61  KLGVVLVGMGNGNGSTLUGVIANREGISWATKDIQQANVFGSLTQASAIRVGSFQGBE 120
      Db      2  RTGILVVGARGSVATTVTYGAALRAGLVPAT-----GCVSALEA----FDG-- 44
OY      121  IYAPFKSLPMVNPDDIVFGGWDISNMNL--ADAMARAKVPDIDLQQLRPYMESMLP- 176
      Db      45  -----VPLRGFDELVFGHDVGTGLVKRAEQLEAGV-----VPR 80
OY      177  -LPGIYDPDFIAANOEBERANNVIKGTKQ-----EQVOQIIKDIAKFEATKVDKXV 226
      Db      81  GLPGVLTAELEDAADAERIP--APGTBEKGDTGTPQASAGAIYADLTGFRERLGLDRYV 138
OY      227  VLTANTERYSNLVVGGLDNTMENTLAAVDRNEAELSPSTLYAIACMENVPRTNGSPQ-N 285
      Db      139  VVNVSTQPPA-VPHPAHSAALREBALRGERPLPVSSLYAYALRAACAFVDFTPSTG 197
OY      286  TFVPGILDLAIARNTLIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGMN 345
      Db      198  ARLPALDELAREQGLPYAGSDGKTGETLVKSVLAMPARALRVRSWSGTNLGGDGAT 257
OY      346  LSAPQTFRSKEISKSNVVDMMVNSNAILYEPGEHPDHVVYIKYVPYVGDSKRAMDEYTS 405
      Db      258  IADPERVVSKNASKGLVLE-----AELGHAVEGGVHHHVDPDLGEWKTAMDHVTFE 308
OY      406  IFMGKSTVLHNTCEDSLAAPITLIDLVLAEISTR 442
      Db      309  CFLGAKMTLQFTWQGDSSLAAPLVLDLARFMALNHR 345
```

Search completed: June 7, 2005, 16:47:40
Job time : 28.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:46:53 ; Search time 81.5 Seconds
(without alignments) 2254.032 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MFENFVSESNVXYTETETI.....NIMRACVGLAPENNMILEYK 510

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1595520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2631	100.0	510	US-10-025-003-12	Sequence 12, Appl
2	2631	100.0	510	US-10-718-952-12	Sequence 12, Appl
3	2624	99.7	510	US-10-025-003-14	Sequence 14, Appl
4	2624	99.7	510	US-10-025-003-16	Sequence 16, Appl
5	2624	99.7	510	US-10-718-952-14	Sequence 14, Appl
6	2624	99.7	510	US-10-718-952-16	Sequence 16, Appl
7	2593	98.6	510	US-10-025-003-2	Sequence 2, Appl
8	2593	98.6	510	US-10-025-003-10	Sequence 10, Appl
9	2593	98.6	510	US-10-718-952-2	Sequence 2, Appl
10	2593	98.6	510	US-10-718-952-10	Sequence 10, Appl
11	2588	98.4	510	US-10-025-003-6	Sequence 6, Appl
12	2588	98.4	510	US-10-718-952-6	Sequence 6, Appl

13	2582.5	98.2	511	US-10-424-599-213009	Sequence 213009,
14	2525	96.0	510	US-10-424-599-154864	Sequence 154864,
15	2431	92.4	505	US-10-442-017-15	Sequence 15, Appl
16	2366	89.9	510	US-10-425-115-231852	Sequence 231852,
17	2366	89.9	516	US-10-425-114-625678	Sequence 625678, A
18	2365	89.9	510	US-10-767-701-46278	Sequence 46278, A
19	2364	89.9	510	US-10-739-930-7635	Sequence 7635, Ap
20	2361	89.7	510	US-10-425-115-206814	Sequence 206814,
21	2355	89.5	542	US-10-425-114-58674	Sequence 58674, A
22	2352	89.4	510	US-10-437-963-173252	Sequence 173252,
23	2346	89.2	510	US-09-921-232-11	Sequence 11, Appl
24	2346	89.2	510	US-09-921-330-11	Sequence 11, Appl
25	2346	89.2	510	US-09-921-339-11	Sequence 11, Appl
26	2337	88.8	510	US-09-727-628-2	Sequence 2, Appl
27	2337	88.8	510	US-10-425-115-231850	Sequence 231850,
28	2337	88.8	510	US-10-425-115-231857	Sequence 231857,
29	2337	88.8	512	US-10-425-114-39621	Sequence 39621, A
30	2337	88.8	512	US-10-425-114-53231	Sequence 53231, A
31	2337	88.8	512	US-10-425-114-66216	Sequence 66216, A
32	2336	88.8	510	US-10-425-115-231853	Sequence 231853,
33	2336	88.8	512	US-10-425-114-46916	Sequence 46916, A
34	2221.5	84.4	509	US-10-437-963-187768	Sequence 187768,
35	2203	83.7	510	US-10-425-115-268216	Sequence 268216,
36	2203	83.7	522	US-10-425-114-62888	Sequence 62888, A
37	2203	83.7	522	US-10-425-114-58895	Sequence 58895, A
38	2203	83.7	529	US-10-425-115-231843	Sequence 231843,
39	2133	81.1	458	US-10-425-115-231845	Sequence 231845,
40	2100	79.8	450	US-10-424-599-154863	Sequence 154863,
41	2032	77.2	431	US-09-876-889-348	Sequence 348, App
42	1616	61.4	558	US-10-128-714-3515	Sequence 3515, Ap
43	1612.5	61.3	534	US-10-128-714-8515	Sequence 8515, Ap
44	1608.5	61.1	534	US-10-128-714-8515	Sequence 8515, Ap
45	1558	59.2	332	US-10-425-114-42702	Sequence 42702, A

ALIGNMENTS

RESULT 1

US-10-025-003-12

Sequence 12, Application US/10025003

Publication No. US20030074685A1

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sedastian, Scott

APPLICANT: Grace, John

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACHARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-C

CURRENT APPLICATION NUMBER: US/10/025, 003

CURRENT FILING DATE: 2002-05-07

PRIOR APPLICATION NUMBER: 06/835,751

PRIOR FILING DATE: APRIL 8, 1997

PRIOR APPLICATION NUMBER: PCT/US98/06822

PRIOR FILING DATE: APRIL 7, 1998

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97

SEQ ID NO 12

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-10-025-003-12

Query Match 100.0%; Score 2631; DB 14; Length 510;

Best Local Similarity 100.0%; Pred. No. 9,8e-218;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFENFVSESNVXYTETETIOSVYNYETTELVAHNRNGTYQWYKPSVNYQFKTNTHVP 60

DB 1 MFENFVSESNVXYTETETIOSVYNYETTELVAHNRNGTYQWYKPSVNYQFKTNTHVP 60

61 KLGWMLVGMGNNSTLTGCVIANREDISWATKXKIQQANVFSLQASAIRVSGFOEE 120


```
Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPWNPDIDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Db 121 IYAPFKSLPWNPDIDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIDIAIARNT 300
Db 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIDIAIARNT 300
Qy 301 LIGGDDPKSGOTKMKSVLVDFVGAGIKPTISVSYNHLGNNDGNMLSAPOTFRSKEISKS 360
Db 301 LIGGDDPKSGOTKMKSVLVDFVGAGIKPTISVSYNHLGNNDGNMLSAPOTFRSKEISKS 360
Qy 361 NVVDMDMNSNAIIEPGEHPDHVVVVKVPYVGDSCRAMDEYTSSEIFMGKNITVLHNTC 420
Db 361 NVVDMDMNSNAIIEPGEHPDHVVVVKVPYVGDSCRAMDEYTSSEIFMGKNITVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480
Db 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 2

```
US-10-718-952-12
; Sequence 12, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-12
```

```
Query Match 100.0%; Score 2631; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 9.8e-218;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFENFKVESPNNVKTETETIOSVYNYETTELVHENRNGTYOMIVKPSVNYQFKTNTHVP 60
Db 1 MFENFKVESPNNVKTETETIOSVYNYETTELVHENRNGTYOMIVKPSVNYQFKTNTHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPWNPDIDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Db 121 IYAPFKSLPWNPDIDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIDIAIARNT 300
Db 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIDIAIARNT 300
```

```
Qy 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIDIAIARNT 300
Db 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIDIAIARNT 300
Qy 301 LIGGDDPKSGOTKMKSVLVDFVGAGIKPTISVSYNHLGNNDGNMLSAPOTFRSKEISKS 360
Db 301 LIGGDDPKSGOTKMKSVLVDFVGAGIKPTISVSYNHLGNNDGNMLSAPOTFRSKEISKS 360
Qy 361 NVVDMDMNSNAIIEPGEHPDHVVVVKVPYVGDSCRAMDEYTSSEIFMGKNITVLHNTC 420
Db 361 NVVDMDMNSNAIIEPGEHPDHVVVVKVPYVGDSCRAMDEYTSSEIFMGKNITVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480
Db 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 3

```
US-10-025-003-14
; Sequence 14, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-14
```

```
Query Match 99.7%; Score 2624; DB 14; Length 510;
Best Local Similarity 99.8%; Pred. No. 3.9e-217;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MFENFKVESPNNVKTETETIOSVYNYETTELVHENRNGTYOMIVKPSVNYQFKTNTHVP 60
Db 1 MFENFKVESPNNVKTETETIOSVYNYETTELVHENRNGTYOMIVKPSVNYQFKTNTHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPWNPDIDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Db 121 IYAPFKSLPWNPDIDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIDIAIARNT 300
Db 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIDIAIARNT 300
```


Db 241 VGLNDTMENTLLAAVDRNBAEISPSSTLYAIACMENVPPINSPOPTFPGGLDLAIANT 300
Qy 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTSTIVSYNHLGNDGNMLSAPOTFRSKEISKS 360
Db 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTSTIVSYNHLGNDGNMLSAPOTFRSKEISKS 360
Qy 361 NVDDMVNSNALIYEPGHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Db 421 EDSLAAPIIIDLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Qy 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510

RESULT 4

US-10-025-003-16
; Sequence 16, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-16

Query Match 99.7%; Score 2624; DB 14; Length 510;
Best Local Similarity 99.8%; Pred. No. 3.9e-217;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MFIEHFKESPNVKTETEIOGVNYETTELVEHNRNGTYQMIYKPKSVNTQFKTNTHVP 60
Db 1 MFIEHFKESPNVKTETEIOGVNYETTELVEHNRNGTYQMIYKPKSVNTQFKTNTHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPWNPPDDIVFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWPLPGI 180
Db 121 IYAPFKSLPWNPPDDIVFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTKEQVOQIIKDIFAKKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEQVOQIIKDIFAKKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENTLLAAVDRNBAEISPSSTLYAIACMENVPPINSPOPTFPGGLDLAIANT 300
Db 241 VGLNDTMENTLLAAVDRNBAEISPSSTLYAIACMENVPPINSPOPTFPGGLDLAIANT 300
Qy 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTSTIVSYNHLGNDGNMLSAPOTFRSKEISKS 360
Db 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTSTIVSYNHLGNDGNMLSAPOTFRSKEISKS 360

Qy 361 NVDDMVNSNALIYEPGHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Db 421 EDSLAAPIIIDLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Qy 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510

RESULT 5

US-10-718-952-14
; Sequence 14, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-14

Query Match 99.7%; Score 2624; DB 16; Length 510;
Best Local Similarity 99.8%; Pred. No. 3.9e-217;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MFIEHFKESPNVKTETEIOGVNYETTELVEHNRNGTYQMIYKPKSVNTQFKTNTHVP 60
Db 1 MFIEHFKESPNVKTETEIOGVNYETTELVEHNRNGTYQMIYKPKSVNTQFKTNTHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPWNPPDDIVFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWPLPGI 180
Db 121 IYAPFKSLPWNPPDDIVFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTKEQVOQIIKDIFAKKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEQVOQIIKDIFAKKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENTLLAAVDRNBAEISPSSTLYAIACMENVPPINSPOPTFPGGLDLAIANT 300
Db 241 VGLNDTMENTLLAAVDRNBAEISPSSTLYAIACMENVPPINSPOPTFPGGLDLAIANT 300
Qy 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTSTIVSYNHLGNDGNMLSAPOTFRSKEISKS 360
Db 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTSTIVSYNHLGNDGNMLSAPOTFRSKEISKS 360
Qy 361 NVDDMVNSNALIYEPGHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480


```

; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-10

Query Match      98.6%; Score 2593; DB 14; Length 510;
Best Local Similarity 98.4%; Pred. No. 1.8e-214;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MFIEFKYSPNVKKTETETIOSVYNYETTELVEHNRNGTYQWIVPKSVNTQKTNTHVP 60
DB 1 MFIEFKYECNPVKKTETETIOSVYNYETTELVEHNRNGTYQWIVPKSVKTEFKTNHVP 60
QY 61 KLGWLVGMGNGNSTLTGVIANREDISMAATKDIQQANFGSLTQASAIRVGSFQGBE 120
DB 61 KLGWLVGMGNGNSTLTGVIANREGISMAATKDIQQANFGSLTQASAIRVGSFQGBE 120
QY 121 IYAPFKSLPMVNPDDIYFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESMPLPGI 180
DB 121 IYAPFKSLPMVNPDDIYFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESMPLPGI 180
QY 121 IYAPFKSLPMVNPDDIYFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESMPLPGI 180
DB 121 IYAPFKSLPMVNPDDIYFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESMPLPGI 180
QY 181 YDPDFIAANQEBRANNVKGTKEQVOQIIKDIAFKKATKVDKVVVMTANTERTYSLV 240
DB 181 YDPDFIAANQEBRANNVKGTKEQVOQIIKDIAFKKATKVDKVVVMTANTERTYSLV 240
QY 241 VGLNDTMENLLAAVDNRNABISPSSTLYAIACMENVPPINSPOPTFPGILDIAIANT 300
DB 241 VGLNDTMENLLAAVDNRNABISPSSTLYAIACMENVPPINSPOPTFPGILDIAIANT 300
QY 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTISIVSYNHLGNNDGNMLSAPOTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTISIVSYNHLGNNDGNMLSAPOTFRSKEISKS 360
QY 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
DB 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
QY 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTTPVN 480
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 9
US-10-718-952-2
; Sequence 2, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
```

```

; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-2

Query Match      98.6%; Score 2593; DB 16; Length 510;
Best Local Similarity 98.4%; Pred. No. 1.8e-214;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MFIEFKYSPNVKKTETETIOSVYNYETTELVEHNRNGTYQWIVPKSVNTQKTNTHVP 60
DB 1 MFIEFKYECNPVKKTETETIOSVYNYETTELVEHNRNGTYQWIVPKSVKTEFKTNHVP 60
QY 61 KLGWLVGMGNGNSTLTGVIANREDISMAATKDIQQANFGSLTQASAIRVGSFQGBE 120
DB 61 KLGWLVGMGNGNSTLTGVIANREGISMAATKDIQQANFGSLTQASAIRVGSFQGBE 120
QY 121 IYAPFKSLPMVNPDDIYFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESMPLPGI 180
DB 121 IYAPFKSLPMVNPDDIYFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESMPLPGI 180
QY 121 IYAPFKSLPMVNPDDIYFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESMPLPGI 180
DB 121 IYAPFKSLPMVNPDDIYFGGWDISNNMLADAMARAKYFDIDLOKQLRPYMESMPLPGI 180
QY 181 YDPDFIAANQEBRANNVKGTKEQVOQIIKDIAFKKATKVDKVVVMTANTERTYSLV 240
DB 181 YDPDFIAANQEBRANNVKGTKEQVOQIIKDIAFKKATKVDKVVVMTANTERTYSLV 240
QY 241 VGLNDTMENLLAAVDNRNABISPSSTLYAIACMENVPPINSPOPTFPGILDIAIANT 300
DB 241 VGLNDTMENLLAAVDNRNABISPSSTLYAIACMENVPPINSPOPTFPGILDIAIANT 300
QY 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTISIVSYNHLGNNDGNMLSAPOTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTISIVSYNHLGNNDGNMLSAPOTFRSKEISKS 360
QY 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
DB 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
QY 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTTPVN 480
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 10
US-10-718-952-10
; Sequence 10, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
```

```

: PRIOR APPLICATION NUMBER: PCT/US98/06832
: PRIOR FILING DATE: APRIL 7, 1998
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Microsoft Office 97
: SEQ ID NO: 10
: LENGTH: 510
: TYPE: prt
: ORGANISM: Glycine max
: US-10-718-952-10
```

Query Match	Score	DB	Length
98.6%	2593	16	510

Qy	1	MFENFKESNNVXYTETETEIQISVYVNTTTELVLHENRGTYQWIVKPSVNVYQFNTINHP	60
Dp	1	MFENFVBECCNVXYTETETEIQISVYVNTTTELVLHENRGTYQWIVKPSVNVYQFNTINHP	60
Qy	61	KLGVMLVGMGNNGSTLTGVIANREDISWATKDXICQANFSGSLTQASAIRGSGFQEE	120
Dp	61	KLGVMLVGMGNNGSTLTGVIANREGISWATKDXICQANFSGSLTQASAIRGSGFQEE	120
Qy	121	IYAPFKSLLPVNVDDIVFGGMDISNNMLADAMARAKVPDIDLOKOLRPYMESWVPLPGI	180
Dp	121	IYAPFKSLLPVNVDDIVFGGMDISNNMLADAMARAKVPDIDLOKOLRPYMESWVPLPGI	180
Qy	181	YDPDFIANOEERANNYIKGTQBOVOQIIXKIDIKAFKATKVDKRVVYLMTANTERYSNLV	240
Dp	181	YDPDFIANOEERANNYIKGTQBOVOQIIXKIDIKAFKATKVDKRVVYLMTANTERYSNLV	240
Qy	241	VGLNDTMENTLAAADRNBAEISPTLYALACWENVPINGSPONTVPPLIDILAIRNT	300
Dp	241	VGLNDTMENTLAAADRNBAEISPTLYALACWENVPINGSPONTVPPLIDILAIRNT	300
Qy	301	LIGGDDFKSGQTKKXSVLVFLVGAGIKPTISVSYNHLGNNDGNMLAPQTFRSKEIISKS	360
Dp	301	LIGGDDFKSGQTKKXSVLVFLVGAGIKPTISVSYNHLGNNDGNMLAPQTFRSKEIISKS	360
Qy	361	NVVDMDNSNALIYERBEHPDHVVYIKVPVPGSGKAMDEYSEIIMGGANTVLHNTC	420
Dp	361	NVVDMDNSNALIYERBEHPDHVVYIKVPVPGSGKAMDEYSEIIMGGASTVLHNTC	420
Qy	421	EDSLAAPITLIDLVLALBELSTRIOFKANECKEKFHSFHPVATILISYLRKAPLVPEGTFRVN	480
Dp	421	EDSLAAPITLIDLVLALBELSTRIEFKANECKEKFHSFHPVATILISYLRKAPLVPEGTFRVN	480
Qy	481	ALSKORAMLENTIRACVGLAPENNMLIEYK 510	
Dp	481	ALSKORAMLENTIRACVGLAPENNMLIEYK 510	

RESULT 11
 US-10-025-003-6
 ; Sequence 6, Application US/10025003
 ; Publication No. US20030074685A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hitz, William
 ; APPLICANT: Sebastian, Scott
 ; APPLICANT: Grace, John
 ; APPLICANT: Streitz, Leon
 ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
 ; FILE OF INVENTION: SACCHARIDES AND PHYTIC ACID
 ; FILE REFERENCE: BB-1077-C
 ; CURRENT APPLICATION NUMBER: US/10/025,003
 ; CURRENT FILING DATE: 2002-05-07
 ; PRIORITY APPLICATION NUMBER: 08/835,751
 ; PRIORITY FILING DATE: APRIL 8, 1997
 ; PRIORITY APPLICATION NUMBER: PCT/US98/06822
 ; PRIORITY FILING DATE: APRIL 7, 1998
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 6
 ; LENGTH: 510

```

; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-6

```

Query Match	98.4%;	Score 2588;	DB 14;	Length 510;
Best Local Similarity	98.2%;	Pred. No. 5e-214;		
Matches 501; Conservative	4;	Mismatches 5;	Indels 0;	Gaps 0;

Qy	1	MFLENKFSYPMVKKTETETESIOGVYVETTELVEHNRGTYOMIKPSPVVOYFQKTNHVP	60
Dp	1	MFLENKJCEBPVKKTETETESIOGVYVETTELVEHNRGTYOMIKPSPVCEYEFKTNHVP	60
Qy	61	KJGVMLVGGANGSGTLTGCVIANREDISMAKTOKIOQANFQSLTQASAIRGSPGEE	120
Dp	61	KJGVMLVGGANGSGTLTGCVIANREGISMAKTOKIOQANFQSLTQASAIRGSPGEE	120
Qy	121	IYAPFSKLLPMYNPDDIYFGCHDISNNMLADAMARAKYFDIDLOKOPRPMESKVPJPGI	180
Dp	121	IYAPFSLLPMYNPDDIYFGCHDISNNMLADAMARAKYFDIDLOKOLRPMESKLPJPGI	180
Qy	181	YBDFITAAOBERANNVIKGTKEBOVOOIIDKIKAFKATYKDVKVVLLMTANTERSNLY	240
Dp	181	YBDFITAAOBERANNVIKGTKEBOVOOIIDKIKAFKATYKDVKVVLLMTANTERSNLY	240
Qy	241	VGGNDTMEMLLAANDRNEAISEPSTLYAIACMVENVPIKNGSPONTFVPGJIDLAIRNT	300
Dp	241	VGGNDTMEMLLAANDRNEAISEPSTLYAIACMVENVPIKNGSPONTFVPGJIDLAIRNT	300
Qy	301	LIGGDDFKSGOTKMSVLYDPLVGAGIKPTSIVSYNHLGNNDGNNLSAPOTFRSEIKS	360
Dp	301	LIGGDDFKSGOTKMSVLYDPLVGAGIKPTSIVSYNHLGNNDGNNLSAPOTFRSEIKS	360
Qy	361	NYVDDMVNSNALIYRGEHPDHVVYIKVVPVGDSCAMDEYTSIEIMGAGNTTVLHNTC	420
Dp	361	NYVDDMVNSNALIYRGEHPDHVVYIKVVPVGDSCAMDEYTSIEIMGAGNTTVLHNTC	420
Qy	421	EDSLAAPITLIDLVLAEILSTRIOFKAENEGKFSHFHPVATILISYLTKAPLVPPEGVN	480
Dp	421	EDSLAAPITLIDLVLAEILSTRIEFKANEKGKFSHFHPVATILISYLTKAPLVPPEGVN	480
Qy	481	ALSKORAMLENTMRACVGLAPENNMTILEYK	510
Dp	481	AUSKORAMLENTMRACVGLAPENNMTILEYK	510

```

RESULT 12
US-10-718-952-6
: Sequence 6, Application US/10718952
: Publication No. US20040128713A1
: GENERAL INFORMATION:
: APPLICANT: Hitz, William
: APPLICANT: Sebastian, Scott
: APPLICANT: Grace, John
: APPLICANT: Strelt, Leon
: TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
: TITLE OF INVENTION: SACCCHARIDES AND PHYTIC ACID
: FILE REFERENCE: BB-1077-C
: CURRENT APPLICATION NUMBER: US/10/718,952
: PRIOR APPLICATION FILING DATE: 2003-11-21
: PRIOR APPLICATION NUMBER: 08/835,751
: PRIOR FILING DATE: APRIL 8, 1997
: PRIOR APPLICATION NUMBER: PCT/US98/06822
: PRIOR FILING DATE: APRIL 7, 1998
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 6
: LENGTH: 510
: TYPE: PRT
: ORGANISM: Glycine max
US-10-718-952-6

Query Match      98.4%; Score 2588; DB 16; Length 510;
Best Local Similarity 98.2%; Pred. No. 5e-214;

```

Matches 501; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 MFIEFKVESPNVKTETETESVNYETTELVEHNRNGTYQMIYVPKSVNTQFKTNTYVP 60
Db 1 MFIEFKVCECPVKTETETESVNYETTELVEHNRNGTYQMIYVPKSVKYEFEKTNIHVP 60
QY 61 KLGVWLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
Db 61 KLGVWLVGMGNGNSTLTGVIANREBISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
QY 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKOLRPYMESVPLPGI 180
Db 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKOLRPYMESVPLPGI 180
QY 181 YDPDFIAANOEBERANNVIKGTQEQVOQIIKDIKAFKATKVDKVVVLTANTERYSNL 240
Db 181 YDPDFIAANOEBERANNVIKGTQEQVOQIIKDIKAFKATKVDKVVVLTANTERYSNL 240
QY 241 VGLNDTMENLLAAVDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDIAIARN 300
Db 241 VGLNDTMENLLAAVDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDIAIARN 300
QY 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISK 360
Db 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISK 360
QY 361 NVVDDMVNSNAIIEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLANT 420
Db 361 NVVDDMVNSNAIIEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLANT 420
QY 421 EDLSLAAPILDLVLLAEISTRIOFKAENEGKTHSFHPVATILSYLTRKAPLVPPTGVV 480
Db 421 EDLSLAAPILDLVLLAEISTRIOFKAENEGKTHSFHPVATILSYLTRKAPLVPPTGVV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 13
US-10-424-599-213009
; Sequence 213009, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 213009
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(511)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.pcp
US-10-424-599-213009

Query Match 98.2%; Score 2582.5; DB 15; Length 511;
Best Local Similarity 98.2%; Pred. No. 1.5e-213;
Matches 502; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
```

```
QY 61 KLGVWLVGMG-GNNGSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFQGE 119
Db 61 KLGVWLVGMGNGNSTLTGVIANREBISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
QY 120 EYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKOLRPYMESVPLPG 179
Db 121 EYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKOLRPYMESVPLPG 180
QY 180 IYDPDFIAANOEBERANNVIKGTQEQVOQIIKDIKAFKATKVDKVVVLTANTERYSNL 239
Db 181 IYDPDFIAANOEBERANNVIKGTQEQVOQIIKDIKAFKATKVDKVVVLTANTERYSNL 240
QY 240 VGLNDTMENLLAAVDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDIAIARN 299
Db 241 VGLNDTMENLLAAVDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDIAIARN 300
QY 300 TLIGGDDFKSGOTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISK 359
Db 301 TLIGGDDFKSGOTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISK 360
QY 360 SNVDDMVNSNAIIEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLANT 419
Db 361 SNVDDMVNSNAIIEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLANT 420
QY 420 CEDSLAAPILDLVLLAEISTRIOFKAENEGKTHSFHPVATILSYLTRKAPLVPPTGVV 479
Db 421 CEDSLAAPILDLVLLAEISTRIOFKAENEGKTHSFHPVATILSYLTRKAPLVPPTGVV 480
QY 480 NALSORAMLENIMRACVGLAPENNMILEYK 510
Db 481 NALSORAMLENIMRACVGLAPENNMILEYK 511
```

```
RESULT 14
US-10-424-599-154864
; Sequence 154864, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154864
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1.pcp
US-10-424-599-154864
```

```
Query Match 96.0%; Score 2525; DB 15; Length 510;
Best Local Similarity 94.7%; Pred. No. 1.3e-208;
Matches 483; Conservative 18; Mismatches 9; Indels 0; Gaps 0;

QY 1 MFIEFKVESPNVKTETETESVNYETTELVEHNRNGTYQMIYVPKSVNTQFKTNTYVP 60
Db 1 MFIEFKVESPNVKTETETESVNYETTELVEHNRNGTYQMIYVPKSVKYEFEKTNIHVP 60
QY 61 KLGVWLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
Db 61 KLGVWLVGMGNGNSTLTGVIANREBISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
QY 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKOLRPYMESVPLPGI 180
Db 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKOLRPYMESVPLPGI 180
```

```
Qy 181 YDPDFIAANOEBERANNVTKGTQEOVOQI IKDIAKFEATKVDKVVVLTANTERYSNLV 240
Db 181 YDPDFIAANOEBERANNVTKGTQEOVOQI IKDIREFKESTKVDKVVVLTANTERYSNLV 240
Qy 241 VGLNDTMENTLAADRNAEISPSSTLYAIACMEVNPPTINGSPONTFVPGILDLAIANT 300
Db 241 VGLNDTMENTLAADRNAEISPSSTLYAIACMEVNPPTINGSPONTFVPGILDLAIKINS 300
Qy 301 LIGGDDFSGQTKMKSVLVDFLVAGAGIKPTSIIVSYNHLGNDGNMLSAPOTFRSKEISKS 360
Db 301 LIGGDDFSGQTKMKSVLVDFLVAGAGIKPTSIIVSYNHLGNDGNMLSAPOTFRSKEISKS 360
Qy 361 NVDDDMVNSNALIYEPGEHPDHVVVKYVPYVGDSCRAMDEYTSIEFMGKNTIYLAHNTC 420
Db 361 NVDDDMVNSNALIYEPGEHPDHVVVKYVPYVGDSCRAMDEYTSIEFMGKNTIYLAHNTC 420
Qy 421 EDSLILAAPIIIDLVLLAELSTRIOFKAENEGKRFHFAVATILSYLTAKAPLVPPGTPVYN 480
Db 421 EDSLILAAPIIIDLVLLAELSTRIOFKAENEGKRFHFAVATILSYLTAKAPLVPPGTPVYN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 15
US-10-442-017-15

```
; Sequence 15, Application US/10442017
; Publication No. US20030215950A1
; GENERAL INFORMATION:
; APPLICANT: Battelle Memorial Institute
; APPLICANT: Dai, Ziyu
; TITLE OF INVENTION: Isolated Polynucleotides and Methods of Promoting a Morphology In
; TITLE OF INVENTION: a Fungus
; FILE REFERENCE: BA4-195
; CURRENT APPLICATION NUMBER: US/10/442,017
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60\382,132
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent version 3.2
; SEQ ID NO 15
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Sesamum indicum
US-10-442-017-15
```

Query Match 92.4%; Score 2431; DB 15; Length 505;

Best Local Similarity 92.3%; Pred. No. 1.7e-200;

Matches 466; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

```
Qy 6 FKVESPNVKTETETIQSYVNYETTELVAHNRNGTYQWIVKPSVNYOKTHTVHPKLGVM 65
Db 1 FKVESPNVKTETETIHSYNYETTELVAHNRNGTYQWIVKPKTYVEKFTDTHVPKLGVM 60
Qy 66 LVGMGNGNSTLTGCVIANREDISWATKDIQOANYFGSLTQASAIRVGSFGCEIYAPF 125
Db 61 LVGMGNGNSTLTGCVIANREGISWATKDKOQOANYFGSLTQASSIRVGSFGCEIYAPF 120
Qy 126 KSLPMPVNPDDIVFGGMDISNNMTLADAMARAKVFDIDLQKOLRPYMESMVPJGIIYDPDF 185
Db 121 KSLPMPVNPDDIVFGGMDISNNMTLADAMGRAKVFDIDLQKOLRPYMEHVPJGIIYDPDF 180
Qy 186 IAAOGEERANNVTKGTQEOVOQI IKDIAKFEATKVDKVVVLTANTERYSNLVVGLND 245
Db 181 IAAOGEERANNVTKGTQEOVOQI IKDMDFEKQNKVDKVVVLTANTERYSNLVVGLND 240
Qy 246 TMENLLAADRNEAEISPSSTLYAIACMEVNPPTINGSPONTFVPGILDLAIANTLIGGD 305
Db 241 TAEISMAEVENREAEISPSSTLYAIACVFNVPPTINGSPONTFVPGILDLAIORNSLIGGD 300
Qy 306 DFKSGQTKMKSVLVDFLVAGAGIKPTSIIVSYNHLGNDGNMLSAPOTFRSKEISKS NVVD 365
```

```
Db 301 DFKSGQTKMKSVLVDFLVAGAGIKPTSIIVSYNHLGNDGNMLSAPOTFRSKEISKS NVVD 360
Qy 366 MVNSNALIYEPGEHPDHVVVKYVPYVGDSCRAMDEYTSIEFMGKNTIYLAHNTCEBSLL 425
Db 361 WVASNGILYEPGEHPDHVVVKYVPYVGDSCRAMDEYTSIEFMGKNTIYLAHNTCEBSLL 420
Qy 426 AAPILDLVLLAELSTRIOFKAENEGKRFHFAVATILSYLTAKAPLVPPGTPVYNALSQ 485
Db 421 AAPILDLVLLAELSTRIOFKAENEGKRFHFAVATILSYLTAKAPLVPPGTPVYNALSQ 480
Qy 486 RAMLENIMRACVGLAPENNMILEYK 510
Db 481 RAMLENIMRACVGLAPENNMILEYK 505
```

Search completed: June 7, 2005, 17:03:08
Job time : 82.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2005, 17:11:57 ; Search time 5315 Seconds
(without alignments)
4649.513 Million cell updates/sec

Title: US-10-718-952-2
Perfect score: 2632
Sequence: 1 MFIEHPKVCPCPVKTKTETETI.....NIMRACVGLAPENMILEYK 510

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10718952/runat_06062005.173400.12846/app_query.faeta_1.1294
-DB=GenEmbl1 -Qfmt=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdt -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ps -NORM=prt -HEADING=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10718952.OCGN_1_17010@runat_06062005.173400.12846 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONELILOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2632	100.0	1739	8	AY038802 Glycine m
2	2627	100.0	1782	6	BD075266 Soybean p
3	2627	99.8	1533	6	BD075269 Soybean p
4	2619	99.5	1791	8	AF293970 Glycine m

5	2477	94.1	1959	8	AB009881 Nicotiana
6	2472	93.9	1931	8	AB032073 Nicotiana
7	2472	93.9	1950	6	E27176 Novel INPS
8	2469	93.8	1959	6	BT013505 Lycopersi
9	2453	93.2	1845	8	AF284065 Sesamum 1
10	2439	92.7	1954	8	BT013759 Lycopersi
11	2392	90.9	1538	8	AF120146 Triticum
12	2392	90.9	1538	8	AF120147 Triticum
13	2392	90.9	1538	8	AF120148 Triticum
14	2382	90.9	1538	8	AF43879 Sueda ma
15	2381	90.5	2053	8	MCU32511
16	2378	90.3	1533	6	C0805008 Sequence
17	2378	90.3	1533	6	AX056743 Sequence
18	2378	90.3	1533	8	AY143904 Arabidops
19	2378	90.3	1863	8	AY054202 Arabidops
20	2378	90.3	1864	8	AY053415 Arabidops
21	2374.5	90.2	1845	8	AY028259 Avicennia
22	2369	90.0	1564	8	AY096554 Arabidops
23	2369	90.0	1863	8	AY065415 Arabidops
24	2368	90.0	1781	6	AX054630 Sequence
25	2368	90.0	1781	8	BNU66307
26	2356	89.5	1533	8	AY323824 Xerophyta
27	2355	89.5	1890	8	ATU30250 Arabidops
28	2352.5	89.4	1567	8	BT001931 Arabidops
29	2352.5	89.4	1837	8	AY085931 Arabidops
30	2352.5	89.4	1870	8	AF372954 Arabidops
31	2349	89.2	1842	8	SPW1P85YM
32	2347	89.2	1914	8	AK103501 Oryza sat
33	2347	89.2	1915	8	AK058750 Oryza sat
34	2344	89.1	1535	6	BD073472 Regulated
35	2344	89.1	1665	6	BD073470 Regulated
36	2344	89.1	1665	8	AF056326 Zea mays
37	2340	88.9	1868	8	AB012107 Oryza sat
38	2338.5	88.8	1921	8	ATU04876 Arabidopsi
39	2337.5	88.8	1978	8	CPTNOL4
40	2336	88.8	1931	6	AR137808 Sequence
41	2336	88.8	1931	6	AR170167 Sequence
42	2335	88.7	1959	6	AB151207 Sequence
43	2333	88.6	1936	8	AB059557 Avena sat
44	2331.5	88.6	1871	8	PV038920 Phaseolus v
45	2328	88.4	2152	8	AF056325 Hordeum v

ALIGNMENTS

RESULT 1	AY038802	1739 bp	mRNA	linear	PLN 16-JUN-2001
LOCUS	Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.				
DEFINITION	AY038802				
ACCESSION	AY038802.1	GI:14764465			
VERSION					
KEYWORDS					
SOURCE	Glycine max (soybean)				
ORGANISM	Glycine max				
REFERENCE	1 (bases 1 to 1739)				
AUTHORS	Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.				
TITLE	Biochemical and Molecular Characterization of a Mutation that Confers a Decreased Raffinosecharide and Phytic Acid Phenotype on Soybean Seeds				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1739)				
AUTHORS	Carlson,T.J. and Hitz,W.D.				
TITLE	Direct Substitution				
JOURNAL	Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402, Wilmington, DE 19880-0402, USA				
FEATURES	Location/Qualifiers				
source	1..1739				
	/organism="Glycine max"				
	/mol_type="mRNA"				

/cultivar="Mye"
/db_xref="taxon:3847"
/dev_stage="developing embryo"
11..1543
/EC_number="5.5.1.4"
/function="cyclizes glucose 6-phosphate to
L-myo-inositol-1-phosphate"
/note="mips-1a"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/db_xref="GI:1476446"
/translation="MPIENPKYECNPVKYETETEIQSYVNYETTELHENRNGTYOMIV
KPKSYKERKTNHIVPKLGVMVGMGNNSTLTGVLANREGISMATKDKIOANVF
GSUTQASAIRVSGFOGEETVAPKSLIPVNPDDIVFGMDISNNLADAMARATVPD
IDLQKOLRFYMESMLPPEIYDPDPLAANOEBRANNVKGTQEOVOOIIKOITAPKE
ATKVDKVVLTANTERYSNLVGLNDTMENTLAAVDRNEAISPSTLVALACWENV
PFINSGPQTFVPGILDIAIARNTLIGDDPKSGOTKMSKLVDFLVGAGIKPTISVS
YVHLNNDMDKNSAPQTFRSKEISKSNVVDVNSNALIYEPGEHDHYVVKYVYV
GDSKRAMDEYETEIFMGKSTIVLHNTGDSILAAPILDVLALIELSTRIFKANE
CKEHSFHPVATILSLITKAPLVPPGTPVNVNLSKORAMLENIMRACVGLAPENNMLE
YK"

ORIGIN

Alignment Scores:

Pred. No.: 3,74e-201 Length: 1739
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-718-952-2 (1-510) x AY038802 (1-1739)

QY 1 MetPheIleGluAsnPhelYsValGluCysProAsnValIleTyThrGluThrGluIle 20
DB 11 ATGTCATCGAGATTTTAAAGTTGAGTGTCCTAAAGTAAAGTACACCGAGCTGAGATT 70
QY 21 GlnSerValTyrAsnThrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 71 CAGTCGGTGTACACATACGAAACACCACTGTTACAGAAACGGAATGACACTAT 130
QY 41 GlnTrpIleValIleProIleSerValIleTyrGluPheIleThrAsnIleHisValPro 60
DB 131 CAGTGGATTGCTCAACCCAAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTT 190
QY 61 LysLeuGlyValMetLeuValGlyTrpGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 191 AATTAGGGGTATATGCTTGTGGGTGGGTGGAACCAACGGCTCAACCGGTGGT 250
QY 81 ValIleIleAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 251 GTTATTGCTTACCGAAGGGCATTTTCATGGGCTTCAAGAGCAAGATTCAACAGCCAT 310
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 311 TACTTTGGCTCCTCCACCAAGCTCAAGCTATCCGAGTTGGGCTTCCAGGAGAGGAA 370
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 371 ACTATAGCCCATTCACAGAGCTGCTTCCAAATGGTTAACCTGACAGACATTTGTTGGG 430
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 431 GGAATGGGAATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGCAAGGTTTGAC 490
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 491 ATCGATTTGCAAGAGAGTTGAGGCTTTCATATGGAATTCATGTTCCATCCCGGAATC 550
QY 181 TyrAspProAspPheIleAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 551 TATGACCCCGAATTTCTGCTGCAACCAAGAGGAGCGTGCACCAACGTCATCAAGGGC 610

QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 611 ACAAGCAAGAGCAAGCTTCAACAAATCATCAAGACATCAAGGCTTTAAGGAACGACCC 670
QY 221 LysValAspLysValValIleuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 671 AAAGTGACAAAGGTGGTTGACTGTGAGCTGCCCAACACAGAGAGGTACAGTAATTTGGTT 730
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaValAspArgAsnGluIleGlu 260
DB 731 GTGGGCTTTAAGACACATGAGATCTCTTGCTGTGTGACAGAAATAGGCTGAG 790
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 791 ATTTCTCTTCCACCTTGATTCATTCCTGCTGTGTATGAGAAATGTTCTTCAATTAT 850
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 851 GGAAGCCCTCAGAACACCTTTGTACAGAGGCTGATTGATCTTGCCATCGCGAGAACT 910
QY 301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 911 TTGATTTGGAGATGACTTCAAGAGTGCTCAGACCAAAATGAAATCTGTGTGTTGAT 970
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 971 TTCCTTGCGGGGCTGCTGATTCAGCCAACTTATAGTCAGTTACACCATCTGGGAAAC 1030
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1031 AATGATGGTATGAATCTTGGGCTCCACAACTTCCGTTCCAGGAATCTCCCAAGAGC 1090
QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnHisPro 380
DB 1091 AACGTTGTGATGAATATGCTCAACAGCAATGCAATGCTCTATAGCCTGGGTAACTCCA 1150
QY 381 AspHisValIleValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1151 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGAGACAGCAAGAACCATGAT 1210
QY 401 GlyTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValIleuHisAsnThrCys 420
DB 1211 GAGTACACTTCAGAGATATTCATGAGGTGGAAGAGACACCATTTGTTGCAACAACATGC 1270
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1271 GAGGATTCCTCTTGTAGCTGCTCTATATCTTGGACTTGCTCTTGTGCTAGCTCAGC 1330
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1331 ACTAGATATGAGTTTAAAGCTAAATAGAGGAAATTCATCATCTCAACCCAGTTGCT 1390
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1391 ACCATCTCAGCTTACCTCCACAGGCTCCTGTGTTCCACCGGTATACCACTGGTGAAAT 1450
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1451 GCATTGTCAAAACACGCGCAATGCTGGAACCAATATGAGGCTTGTTGTGATTTGGCC 1510
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1511 CAGAGATATACATGATTTCTCAGATACAG 1540

RESULT 2
BD075266 1782 bp DNA linear PART 27-AUG-2002
LOCUS
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION
BD075266
BD075266.1 GI:22620869
VERSION
JP 2001519665-A/1.
KEYWORDS

SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1782)
AUTHORS Hitz, W.D. and Sebastian, S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose
JOURNAL Patent: JP 2001519665-A 1 23-OCT-2001;
E1 DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR13
PN JP 2001519665-A/1
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC Saccharides
CC and phytic acid
FH Key
FT CDS Location/Qualifiers
1..1782 Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:
Pred. No.: 3,85e-201 Length: 1782
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Dbs: 6 Gaps: 0

US-10-718-952-2 (1-510) x BD075266 (1-1782)

QY 1 MetPheIIeGluAsnPhelYsValIGluCyvProAsnValIyEThrGluThrGluIle 20
Db 54 ATGTTTCATGAGAAATTTTAAAGTTGAGTGTCTTAATGTAGAGTACCAGACTGAGATT 113
QY 21 GlnSerValIyYAsnTyrgIuThrThrgIuLeuValIhIGluAsnArgAnGIyThrTy 40
Db 114 CAGTCGCTTACAACTACGAAACCGAACTTTGTTCAAGAGAACGAAATGCGACTAT 173
QY 41 GlnTPILeValIyBProIySerValIyTyrgIuPhelYsThrAsnIIehIsvaIPro 60
Db 174 CAGTGGATTGTCAAAACCAATCTGTCAATACGAATTTAAACCAATCCATGTTCTT 233
QY 61 LysLeuGluValMetLeuValIGluTyrgIyAsnAsnGlySerThrLeuThrgIyGly 80
Db 234 AAATTAAGGGGTAATGTTGTGGGTGGGGTGGGAAACAGGCTCAACCTCAACGGTGGT 293
QY 81 ValIIeAlaAsnArgGluGlyIleSerTPALaThrIyAspIyEIIeGlnGlnAlaAsn 100
Db 294 GTTATTGCTAACCGAGAGGCAATTCATGCGGCTACAAAGACAAAGATTCAACAGCCAT 353
QY 101 TyrPheGlySerLeuThrgIuAlaSerAlaIIeArgValIGlySerPheGlnGlyGlu 120
Db 354 TACTTTGGCTCCCTCCACCAAGCTCAAGCTATCCGAGTGGGCTCTTCCAGGAGAGGAA 413
QY 121 IleTyraIaProPheYsSerLeuEuprometValIAsnProAspAspIIeValIpheGly 140
Db 414 ATCTAAGCCCATTCAGAGCGCTTCCAAATGATTAAACCTGACGACATGTTGTTGG 473
QY 141 GIlyTPAspIIeSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIyValIpheArg 160
Db 474 GGATGGGATATACGAAACATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 533

QY 161 IleAspLeuGlnIyGlnLeuAsnProTyMetGlySerMetLeuProLeuProGlyIle 180
Db 534 ATCGATTTCAGAACAGATTGAGGCTTTACATGAAATCATGCTTCACCTCCCGGAATC 593
QY 181 TyrAspProAspPheIIeAlaIaAsnGlnGluArgAlaAsnAsnValIleYsGly 200
Db 594 TATGACCCGGATTTCTATCTCTCCCAACAAAGAGGCGTCCAAACACTCATCAAGGCG 653
QY 201 ThrIySGlnGluGlnValIGlnGlnIleIyAspIleIyAlaIpheYsGluAlaThr 220
Db 654 ACAAAAGCAAGACAGTTCAACAAATCATCAAAAGCATCAAGCGCTTTAAGCAAGCCACC 713
QY 221 LysValaAspIyValIyLeuThrPheIaAsnThrgIuArgTySerAsnLeuVal 240
Db 714 AAAGTGAACAGGATGTTGACTGTGAGCTGCACACAGAGGATGACGTAATTTGGTT 773
QY 241 ValIGlyLeuAsnAspThrMetGluAsnLeuValaIaValaAspArgAnGluIaGlu 260
Db 774 GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
QY 261 IleSerProSerThrLeuTyraIIeAlaCyvValMetGluAsnValProPheIIeAsn 280
Db 834 ATTTCTCTTCCACCTTGATGCCATTGGCTGTGTAAGAAATGTTCTTTCATTAAT 893
QY 281 GlysSerProGlnAsnThrPheValaProGlyLeuIIeAspLeuAlaIIeAlaArgAnThr 300
Db 894 GGAACCCCTCAGAACACTTTGTACAGGCGTGATTTGATCTTGCCATCGCGAGAACACT 953
QY 301 LeuIIeGlyGlyAspAspPheYsSerGlyInThrIyMetIySerValIeValaIAsp 320
Db 954 TTGATTGGTGGAGATGACTTCAAGAGTGTCCAGACCAAAATGAAATCTGTGTGTTGAT 1013
QY 321 PheLeuValIGlyAlaGlyIleIyBProThrSerIIeValSerTyraAsnIIeGluIyAsn 340
Db 1014 TTCCTTGTGGGGCTGGATATCAAGCCAACTATATATGATCAAGTTACATCTGGGAAAC 1073
QY 341 AsnAspGlyMetAsnLeuSerIaProGlnThrPheArgSerIyGlnIleSerIySsr 360
Db 1074 AATGATGGTATGAATCTTGGGCTCACAACTTTCCTTCAGAAATCTTCAGAGAC 1133
QY 361 AsnValaIaAspAspMetValaAsnSerAsnAlaIIeLeuTyrgIuProGlyGluIhIAsp 380
Db 1134 AACGTTGTGATGATATGATGTCACAGCAATGCCAATCTTAAGACCTGTGTAACATCA 1193
QY 381 AspHIsvalaIaIleIySryValaProTyraIaGlyAspSerIyArgAlaMetAsp 400
Db 1194 GACCATGTTGTTTATTATTAAGTATGCTTACGTAGGGGACAGCAAGAGACCATGGAT 1253
QY 401 GIlyTyThrSerGluIlePheMetGlyGlyIySsrThrIIeValLeuHIsAsnThrCy 420
Db 1254 GAGTACACTTCAGAAATATTCATGGGTGAAAGACACCAATGTTTGCACAACATGTC 1313
QY 421 GluAspSerLeuLeuAlaIaProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
Db 1314 GAGGATTCCTCTTACGCTCTCTATTAATCTTGAACCTTGCTCTTCTGTACGCTCAGC 1373
QY 441 ThrArgIIeGluPhelYsAlaGluAsnGluGlyIySphenIIsSerPheHIsProValaIa 460
Db 1374 ACTAAGATTCAGATTTAAAGCTGAATAAGGAAATTCACATCTTCCACCAAGTTGCT 1433
QY 461 ThrIleLeuSerTyLeuThrIyAlaProLeuValaProProGlyThrProValaIleAsn 480
Db 1434 ACCATCTTCAGTACCTCACCAAGGCTCTCTGTGTTCCACCGGATACCAATGTGTGAT 1493
QY 481 AlaLeuSerIyGluArgAlaMetLeuGluAsnIleMetArgAlaCyevaIaGlyLeuAla 500
Db 1494 GCATTGTCAAGACAGCTGCATATGCTGAAAAACAAATGAGGGCTTGTGTTGATTGGCC 1553
QY 501 ProGluAsnAsnMetIleuGluTyryIyS 510
Db 1554 CCAGAGATTAACATGATTTCTGAGTACAAAG 1583

RESULT 3

BD075269 1533 bp DNA linear PAT 27-AUG-2002
LOCUS BD075269
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION BD075269
VERSION BD075269.1 GI:22620872
KEYWORDS JP 2001519665-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1533)
AUTHORS Hitz,W.D. and Sebastian,S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid
JOURNAL Patent: JP 2001519665-A 4 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR33
PN JP 2001519665-A/4
PD 23-OCT-2001
PF 07-APR-1998 JP 1298543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ,SCOTT ANTHONY SEBASTIAN
PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC and phytic acid
FH Key Location/Qualifiers
FT CDS 1..1533.
Location/Qualifiers
1..1533
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.:	8 05e-201	Length:	1533
Score:	2627.00	Matches:	509
Percent Similarity:	99.80%	Conservative:	0
Best Local Similarity:	99.80%	Mismatches:	1
Query Match:	99.81%	Indels:	0
DB:	6	Gaps:	0

US-10-718-952-2 (1-510) x BD075269 (1-1533)

QY 1 MetPheIIeGluAsnBheLySValGIuCySProAsnValIysTyrThrgIuThrgIuIle 20
DB 1 ATGTTCAATCGAGATTTTAAAGGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
QY 21 GlnSerValITyrAsnTyrGluThrThrgIuLeuValHieGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCCGCTGTACAACTACCAACACCGAAGCTGTTCACAGAGAACGAGATGGCACTTAT 120
QY 41 GlnTrpIleValIysProIysSerValIysTyrGluPheIysThrsnIleHISValPro 60
DB 121 CAGTGAATGTTCAAACCAAAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTT 180
QY 61 LysLeuGlyValMetLeuValGIuTyrGIuGlyAsnAsnGlySerThrLeuThrgIuGly 80
DB 181 AATTATGGGGTATATGTTGTGGGTTGGGATGGAACACAGGGCTCAACCTCACCGGTGCT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTTPAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATATGCTAACCGAGAGGCATTTCATGGGCTACAAAGACAAAGATTCAACAGCCAAAT 300
QY 101 TyrPheGlySerLeuThrgIuAlaSerAlaIleArgValGIuSerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTACCCAAAGCTCAGCTATCCAGTTGGGCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140

DB 361 ATCTATGCCCATTCACAAAGCGCTGCTTCAATGGTTAAACCTTCAGCATTTGTGTTGGG 420
QY 141 GIYTPAPPIIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGATATCAGCAACATGAACCTGGCTATGTGACATGAGGCGCAAGAGGTTTATAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTGCAGAAACACATGTAGAGGCTTACATGGAATCCATGCTTCCCTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluAlaAsnValIleLysGly 200
DB 541 TATGACCCGGATTTTATGCTGCCAACAGAGGAGCGCTGCCAACAGCATCATCAAGGCG 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAAGCAAGAGCAAGTTCACAAATCATCAAGCATCAAGCGCTTTAAGGAACCCAC 660
QY 221 LysValAspLysValValLeuTyrThrAlaAsnThrgIuArgTyrSerAsnLeuVal 240
DB 661 AAAGTGCACAAAGTGTGTTGTACTGTGACCTGCAACACAGAGGTACAGTATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTAAATGACACCATGAGAACTCTTGGCTGTGGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCACCTTGATATGCAATGCTTGTGTGTATGAAATGTTCTTTCTTATAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAACCCCTCAGAACACTTTTGTACAGAGGCTGATGTATCTTGCCATCGCGAGAACACT 900
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGGAGATGACCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGGTTAT 960
QY 321 PheLeuValGIuAlaGlyIleLysProThrSerIleValSerTyrAsnHieLeuGIuAsn 340
DB 961 TTCTTGTGGGGGCTGGTATCAGCAACATCTATAGCACTTACACACATCTGGGAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGGATGAATCTTTCGGCTCCACAACTTTCGGTCCAGGAATGCTCCAAAGGC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHISPro 380
DB 1081 AACGTTGTTGATGATATGATGTCACACAGCAATGCCATCTTATGAGCCTGGTGAACATCA 1140
QY 381 AspHISValValValIleLysTyrValProTyrValGIuAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTGTGTATTAATGATATGTGCTTACGTAGGGAGACGAATAGACCTGAT 1200
QY 401 GluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValIleuHISAsnThrCys 420
DB 1201 GAGTACACTTCAGAAATATTCATGGGTGGAAGACACACATTTGTTTGCACAAACAGC 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTACGCTCTCTATTAATCTTGGACTTGGCTCTTGGCTGAGCTCAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGlyLysPheHISerPheHISProValAla 460
DB 1321 ACTAGAAATGAGTTTAAAGCTGAATAATAGGAAATATTCATCTACTTCCACCCAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValAlaAsn 480
DB 1381 ACCATCTCAGGTACTACCAAGGCTCTCTGTTCCACGGGATACCAAGTGTGAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGIuLeuAla 500

Db 1441 GCATGTCAAGACGCGTGCATGCTGAAACATATAGAGCGCTTGTTGGATTGGCC 1500

QY 501 PROGLuAsnAsnMetIleLeuGluTyrIvs 510
|||||
Db 1501 CCAGAGAAATPAACATGATTCGAGTACCAAG 1530

RESULT 4
AF293970
LOCUS AF293970
DEFINITION Glycine max myo-inositol-3-phosphate synthase (MIPS) mRNA, complete cds.
ACCESSION AF293970
VERSION AF293970.1 GI:13936690
KEYWORDS
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 1791)
Hegeman,C.E., Good,L.L. and Grabau,E.A.
Expression of D-myo-inositol-3-phosphate synthase in soybean.
Implications for phytic acid biosynthesis
JOURNAL Plant Physiol. 125 (4), 1941-1948 (2001)
MEDLINE 21199082
PUBMED 11299372
REFERENCE 2 (bases 1 to 1791)
Hegeman,C.E., Good,L.L. and Grabau,E.A.
Direct Submission
AUTHORS Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed
JOURNAL Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA
24060, USA

FEATURES
source
1.1791
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/culivar="Williams 82"
/db_xref="taxon:3847"
/dev_stage="immature cotyledons"
1.1791
/gene="MIPS"
62..1594
/gene="MIPS"
/note="isomerase"
/codon_start=1
/product="myo-inositol-3-phosphate synthase"
/protein_id="AAK49896.1"
/db_xref="GI:13936691"
/translation="MTENFKYECNPKVKTTEIEISVNYETTELVEHNRNGTYQNTV
KRSKYVERKTNIHVKLGVMVWGNGNSTLTGVIANRSGISWATKDKIQANVYF
GSLTQASAIRVGSFOGEIYAPFKSLPMVNPDIYFGMDISNMNLADAMARAKVFD
ILOKQMRPYMESMLPLPGIYDPDPIAQCERANNVIGTKOEOVOQIIKDKAKE
ATKVKVVLVMTANTERSYNIIVGNDINMENLLAADREARISPTLYAIAQWENV
PRINSRQTPYRGLIDLAIRANTLIGDDPKSGQTKMSYLVDFVGAGIKPTSVS
VYHLKNDKMNLSAPOTFRSKISKNVDDVNSVALIYEGEHPDHVAVIKYEVY
GUSKRAMDEYTSIEFMGKSTIVLANTCDSLLAAPIIDLVLMELSTRIRPKENE
GKFSHPHVAITLISYTKAPLVPPGTGPNALSKORAMENIMRACVGAAPNNMLLE
YK"

ORIGIN
Alignment Scores:
Pred. No.: 4,266-200 Length: 1791
Score: 2619.00 Matches: 508
Percent Similarity: 99.61% Conservative: 0
Best Local Similarity: 99.61% Mismatches: 2
Query Match: 99.51% Indels: 0
DB: 8 Gaps: 0
US-10-718-952-2 (1-510) x AF293970 (1-1791)

QY 1 MetPheIleGluAsnPheIvsValGluCysProAsnValIleTyrThrGluThrGluIle 20
|||||
|||||

Db 62 ATGTCATGAGAAATTTTAAGGTGAGTGTCTTAATGTGAAGTACACCGAGCATGAGATT 121

QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
|||||
Db 122 CAGTCGGGTAAACAACATCAACAAACCAACGAACTTGTTCAAGAGAACGGAATGCGACTAT 181

QY 41 GlnTPRILEValIysPheIysSerValIysTyrGluPheIysThrAsnIleHisValPro 60
|||||
Db 182 CAGTGGATTTGTCNAACCAAAATCTGTCAAAATCGAAATTTAAACCAACATCCATGTTCT 241

QY 61 LysLeuGlyValMetLeuValGlyTyrGlyIysAsnAsnGlySerThrLeuThrGlyGly 80
242 AATTAGGGGTATAGCTTTGGGTTGGGTGGGAACAAGGCTCAACCTCAGCGGTGGT 301

QY 81 ValIleAlaAsnArgGluGlyIleSerTPRAlaThrLysAspLysIleGlnGlnAlaAsn 100
302 GTTATTGCTAACACCGAGAGGCAATTCATGGGCTCAAAAGACAAGATTCAACCAACGAAAT 361

QY 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGlu 120
362 TACTTGGCTCCCTCACACCAAGCTTCAGCTATCCAGTTGGGTCTTCCAGGAGAGAA 421

QY 121 IleTyrAlaProPheIysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
422 ATCTAAGCCCATTCACAGAGCTGCTTCCAAATGGTTAACCTGACGACATTTGTTGGG 481

QY 141 GlyTPRAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
482 GGATGGGATATCAGCAACATGAACCTGCTGATGTCATGCCAGGCAAGGTTGTTGAC 541

QY 161 IleAspLeuGlnIysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
542 ATCGATTTCAGAAAGAGAGGAGGCTTACATGAAATCCATCTTCACTCCCGGAATC 601

QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluIysArgAlaAsnAsnValIleLysGly 200
602 TATGACCCGGATTTCTATCTCTCCCAACCAAGAGAGCGTCCAAACAGCTCATCAAGGCG 661

QY 201 ThrLysGlnGluIysAlaGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
662 ACAAGACAGAACCAAGTTCACAAATCATCAAGACATCAAGCGCTTTAAGAACGACACC 721

QY 221 LysValAspLysValValValLeuThrPheAlaAsnThrGluIysTyrSerLeuLeuVal 240
722 AAAGTGAACAAGTGTGTACTGTGTGAGATGTGCAACACAGAGGTACGTATTTGTT 781

QY 241 ValGlyLysAsnAspThrMetGluAsnLeuLeuAlaAlaValAspArgAsnGluIys 260
782 GTGGGCTTAATGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGCTGAG 841

QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
842 ATTTGCTCTTCAACCTTGATGCCATGTGTTGTTATGGAAGAAAGTTCTCTTCAATTAAT 901

QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
902 GGAAGCCCTCAGAAACATTTTATCACAGGCGATTTGAATCTTGCCATCCCGAGAACACT 961

QY 301 LeuIleGlyLysAspAspPheLysSerGlyIleThrLysMetLysSerValLeuValAsp 320
962 TTGATTGGTGGAGATGATTTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTTGGTGAAT 1021

QY 321 PheLeuValGlyAlaGlyIleLysPheProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
1022 TTCCTTGGGGGCTGGTATCAAGCAACATTAATGTCAGTTACCAACATCGGGAAGAAC 1081

QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
1082 AATGATGTATGAATCTTTCGGCTCCACAAATCTTCCTCCAAAGAAATCTTCCAAAGAAC 1141

QY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
1142 AACGTTGTTGATGATGTGTCAACAGCAATGCAATCCTTAATGAGCTGTGTAAATCTCA 1201

QY 381 AepHisValIValIleYsTYrValProTYrValGlyAspSerLYsArgAlaMetAsp 400
Db 1202 GACCAAGTGTGTATTATTAAGATGTGCTTACGTAGGGGACAGACAGACCAATGAT 1261
QY 401 GltYrTYrThSerGluIlePheMetGlyGlyLYsSerThIleValIleuHisAnthCYs 420
Db 1262 GAGTACACTTCAGACATATTCATGCGTGAAGACACCATTTGTCACAACATCTGC 1321
QY 421 GluAspSerLeuLeuAlaIleProIleIleuAspLeuValIleuAlaGluLeuSer 440
Db 1322 GAGGATTCCTCTGAGCTCTCTCTATTATCTTGAGATGTGCTCTCTTCTGAGCTACAC 1381
QY 441 ThrArgIleGluPheLYsAlaGluLeuGluGlyLYsPheHisSerPheHisProValAla 460
Db 1382 ACTAGATTCGAGTTTAACTGAAATGAGGAAAATTCACATCTCCACCCAGTTGCT 1441
QY 461 ThrIleLeuSerTYrLeuThrLYsAlaProLeuValProProGlyThrProValValAsn 480
Db 1442 ACCATCTACGCTACCTCACCAAGGCTCTCTGTTCCACCGGTCACCGGTCGTAAT 1501
QY 481 AlaLeuSerLYsGluArgAlaMetLeuGluAsnIleMetArgAlaCYsValGlyLeuAla 500
Db 1502 GCAATGTCAAAGCAGCTCAATGCTGGAACATATATGAGGCTTGTGTGATTTGGCC 1561
QY 501 ProGluAsnAsnMetIleLeuGluTYrLYs 510
Db 1562 CCAGAGATTAATCATGATTTCTCGAGTACAG 1591
RESULT 5
AB009881 1959 bp mRNA linear PLN 13-JUN-2000
LOCUS Nicotiana tabacum mRNA for myo-inositol 1-phosphate synthase,
DEFINITION complete cds.
ACCESSION AB009881.1 GI:8096265
VERSION myo-inositol 1-phosphate synthase.
KEYWORDS Nicotiana tabacum (common tobacco)
SOURCE Nicotiana tabacum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1 (sites)
REFERENCE Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
AUTHORS Screening of wound-responsive genes identifies an immediate-early
TITLE expressed gene encoding a highly charged protein in mechanically
wounded tobacco plants
JOURNAL Plant Cell Physiol. 41 (6), 684-691 (2000)
MEDLINE 20399434
PUBMED 10945337
REFERENCE 2 (bases 1 to 1959)
AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1997) Kojiro Hara, Nara Institute of Science and
Technology, Plant Molecular Breeding, Takayama 891e-5, Ikoma, Nara
630-0101, Japan (E-mail:k-hara@nbs.aisr-nara.ac.jp,
Tel:+81-743-72-5653 (ex.5653), Fax:+81-743-72-5659)
FEATURES
source Location/Qualifiers
1..1959
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/db_xref="taxon:4097"
/cissue_type="leaf"
94..1626
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="BA95788.1"
/db_xref="GI:8096266"
/translation="MPIENPKYSESPIVKATSEIHSIVYDTTELVEHEKNGYQMTV
KRTVYERKTDVHPKIGVIMVGGNGSGTLTGIVLINREGISWATKDKYQANVF
GSITQASTIRVSGFNEEYIAPFKSLIPVNPDDVYFGMDISGNLADAMARAYED
IDIQOLRPMESNVPFGIYIPDFIAAQSGRANNVITGTRKEQIDQIIRKREFKE
KKIKVKVVLWANTERYSNVVGVLDITMENLFASVDREAEISPTLYAIACILENV

PFINGSPONTFVPGILIDIAIKRNTLIGDDFKSGQTKKSVLVDFLNGAIKPKPSIVS
YNHIGNDGNMLSAIPOTERSKEISKNVNDVMSNMLLYEPGHPIHVAVIKVYPV
GDSKRAMEYSEIFMGKNITIVLANTLEIPIIDILVLAELSTRIOAKGE
GKPSFPHVATILSYLFKAPLVPPGTPIVNVNLSKORAMLENTILACVGLABENNMILE
YK"
1959
/note="15 a nucleotides"
ORIGIN
polyA_site
Alignment Scores:
Pred. No.: 1,1e-188 Length: 1959
Score: 2477.00 Matches: 472
Percent Similarity: 96.67% Conservative: 21
Best Local Similarity: 92.55% Mismatches: 17
Query Match: 94.11% Indels: 0
Gaps: 0
DB: 8
US-10-718-952-2 (1-510) x AB009881 (1-1959)
QY 1 MetPheIleGluAsnPheLYsValGlyCYsProAsnValLYsTYrThGluThrGluIle 20
Db 94 ATGTTATTAGGAACCTTTAAGTTGAGAGCCCAACGTTAAGTACACCGAAAGTGAGATT 153
QY 21 GlnSerValTYrAsnTYrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTYr 40
Db 154 CACTCTGCTATGATTTATCAACCACTAGATTGTTCAATGAGGAGAAAATGGAGCTTAC 213
QY 41 GlnTPriLeValLYsProLYsSerValLYsTYrGluPheLYsThrAsnIleHisValPro 60
Db 214 CAATGCACTGCACACCTTAAGACTGTCAATATAGATTCAAGACTGATTCATTTCC 273
QY 61 LysLeuGlyValMetLeuValGlyTYrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 274 AATATTAGGGTTATATCTTGTGATGGGTTGGAACATGTTCAACTTGACCGGTGCT 333
QY 81 ValIleAlaAsnArgGluGlyIleSerThrPalaThrLYsAspLYsIleGlnGlnAlaAsn 100
Db 334 GTTATTGCTAACAGAGAAAGAAATTTCAATGCGCCACCAAGATTAAGTGCACACAGCCAAAT 393
QY 101 TYrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
Db 394 TACCTTGGCTCTTACCTACGAGCTTCTACTATTCAGATTTGGGCTCTTTCATATGAGAAAG 453
QY 121 IleTYrAlaProPheLYsSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 454 ATCTATGCTCAATTTAAAGCTCTCTTCAATGTCATCCAGATGAGTATGTTGGA 513
QY 141 GltYrPheAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLYsValPheAsp 160
Db 514 CGATGGACATCAGCGGATGAAATTTGCAGATGCCATGGCAAGGGCTTAAGGATTTGAT 573
QY 161 IleAspLeuGlnLYsGluLeuArgProTYrMetGluSerMetLeuProLeuProGlyIle 180
Db 574 ATTTGATCTCAAAACACAGTTGAGGCCCTTACATGAAATTTATGTCTCCACTACTCTGTATC 633
QY 181 TYrAspProAspPheIleAlaIleAsnGlnGluGlyIleArgAlaAsnAsnValIleLYsGly 200
Db 634 TATGATTCGATTTCTATTGCTGCTAACCAAGGTCACGCGCCAAACACGTGATCAAAGGA 693
QY 201 ThrLYsGlnGluGlnValGlnGlnIleIleLYsAspIleLYsAlaPheLYsGluAlaThr 220
Db 694 ACCAAGAAAGAAACAAATGATCAATTCATTAAGGATTTAAGGACTTTAAGAGAAAGAAC 753
QY 221 LysValAspLYsValValIleuThrThrAlaAsnThrGluArgTYrSerAsnLeuVal 240
Db 754 AAAGTGAACAAGCTGTATCTTGTGACCTGCTTACCTGAAGATACAGTAATGTGGTT 813
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluIleGlu 260
Db 814 GTTGACCTTAATGATACATGAAATCTCTTGTCTGTGGAACAGAAATGAAGTAACTGAA 873
QY 261 IleSerProSerThrLeuTYrAlaIleAlaCYsValMetGluAsnValProPheIleAsn 280

Db 874 ATATCTCTTCCACTTGTATGATGATTCGTCGATTCCTGAAAATGTCCTTTATTAAT 933
 Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaGAsnThr 300
 Db 934 GGAAAGCCCAAGAACCTTTGTGCCAGGCTCATGATTTGGCCATCAAGAGAACACA 993
 Qy 301 LeuIleGlyValAspAspPheLeuSerGlyGlnThrIleMetLeuSerValIleValAsp 320
 Db 994 CTGATGCTGTGTATGATGATCTTTAAGAGTGTCAACCAAAATGAGTGTGCTGTGAT 1053
 Qy 321 PheLeuValGlyValGlyIleLeuProThrSerIleValSerTyrAsnIleGlyAsn 340
 Db 1054 TTCCTTGTGGAGCTGGATTAAGCCAAATGATGATGCTCAACCACTTTGGGTAC 1113
 Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGlyIleSerLeuSer 360
 Db 1114 AAGAGAGATGATGATGCTGTGCTCCCAAACTTCGCTCAAAAGAGATCTCAAAAGT 1173
 Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlyIleProGlyGlyIlePro 380
 Db 1174 AATGTTGTGATGATGATGCTTTCAAGCAATGCAATCTTTATAGCTGGAGACACCT 1233
 Qy 381 AspHisValValValIleLeuTyrValProTyrValGlyAspSerIleAspAlaMetAsp 400
 Db 1234 GACCATGTTGTTGATTAAGTATGTCATATGTGGAGACAGCAAAAGGCAATGAT 1293
 Qy 401 GlyTyrThrSerGlyIlePheMetGlyGlyLeuSerThrIleValLeuHisAsnThrCys 420
 Db 1294 GAGTACACATCTGAGATTTTCATGCGAGGAGAAACACCATGTTTGGCAACAATACCTGT 1353
 Qy 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGlyLeuSer 440
 Db 1354 GAGGATCTCTTTTGGCTGCTCCAAATTAATCTTGATTTGCTCTTCTTGTAACCTAGT 1413
 Qy 441 ThrArgIleGluPheLeuValGluAsnGlyIlePheHisSerPheHisProValAla 460
 Db 1414 ACCCGCATTCAGCTCAAGCTGAAGGAGGATTAAGTTCACCTTCCACCCCGGTGCT 1473
 Qy 461 ThrIleLeuSerTyrLeuThrIleValAlaProLeuValProProGlyThrProValIleAsn 480
 Db 1474 ACTATCTCAAGCTTACCTTACCAAGGCTCTCTGTACACCAAGTACACCAAGTGTGAAT 1533
 Qy 481 AlaLeuSerIleGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
 Db 1534 GCACTCTCAAGAGAGGAGCAATGCTTTGAGAACATATGAGGGCTGTGTGSACTTGA 1593
 Qy 501 ProGluAsnAsnMetIleLeuGlyTyrIle 510
 Db 1594 CCAGAGAACACATGATTCGGAATCAAA 1623
 RESULT 6
 AB032073 1931 bp mRNA linear PLN 08-SEP-1999
 LOCUS Nicotiana paniculata NpINP51 mRNA for myo-inositol-1-phosphate
 DEFINITION synthase, complete cds.
 ACCESSION AB032073.1 GI:5834499
 VERSION AB032073.1
 KEYWORDS myo-inositol-1-phosphate synthase.
 SOURCE Nicotiana paniculata
 ORGANISM Nicotiana paniculata
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Nicotiana.
 REFERENCE
 AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
 TITLE Myo-inositol-1-phosphate synthase
 JOURNAL Published Only in Database (1999)
 AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Akiko Hashimoto, Japan Tobacco Inc., Plant
 Breeding and Genetic Research Laboratory; 700 Higashibara, Iwate,
 Toyoda-cho, Shizuoka 438-0802, Japan

(E-mail: Akiko.Hashimoto@pbgrl.jti.co.jp, Tel.: +81-538-32-7116,
 Fax: +81-538-33-6046)
 FEATURES
 source
 location/Qualifiers
 1. 1931
 /organism="Nicotiana paniculata"
 /mol_type="mRNA"
 /db_xref="taxon:62141"
 gene
 1. 1931
 /gene="NpINP51"
 CDS
 92..1624
 /gene="NpINP51"
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase"
 /protein_id="BA84084.1"
 /db_xref="GI:5834500"
 /translation="MPIENFKVESPVNVKTESEHSVVDYQTTELVDHDKNGYQWTV
 KPTVYVEKRTVHVVKLGVMVWGMNGNSTLTGVIANREGISVATKDKVQAAVYF
 GSLTQASTIRVGSFNGEEIYAPPEKSLPVPNPDDVYVGGMDISDKMLADAMAKYFD
 IDLQKLRPYMESMPVPGIYDPDFIAAGSRANVITGKKEQIDQIYKQIREKE
 KKKVDVVLTMTANTRYSNVVYGLDNTMENLPAFSDRBAEISBSTVAIACILENV
 PTNPSQKQTFVPPGLDLAIKKNVILIGSDPFGSKYKMSVLVDPLVAGIKPTLSVS
 YNHLGNNDQMLNAPQTFRSKEISKNVVDWVSNAILYEPGEHDPHVIVIKVPPV
 GDSKRAMDEVTSEIFWGGKNTIVLHNTCEDSLIAPILIDLVLALSTRIQKAGE
 GKRFSPHVPATILSYLTAKAPLVPGTPVNVNLSKQRAMLENILIRACVGLAPENNMLE
 YK"

ORIGIN

Alignment Scores:
 Pred. No.: 2,72e-188 Length: 1931
 Score: 2472.00 Matches: 470
 Percent Similarity: 96.86% Conservative: 24
 Basic Local Similarity: 92.16% Mismatches: 16
 Query Match: 93.92% Indels: 0
 DB: 8 Gaps: 0

US-10-718-952-2 (1-510) x AB032073 (1-1931)

Qy 1 MetPheIleGluAsnPheLeuValGlyCysProAsnValIleGlyThrGluThrGluIle 20
 Db 92 ATGTTATTTAGAGAAATTTTAAAGTTGAGACCCCAAGTTAAAGACCCAAATGAAATT 151
 Qy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyTyrTyr 40
 Db 152 CACTCTGTCTATGATTAATCAACACCTGAGTTAGTTCAATGATGAGAAAATGGACATAT 211
 Qy 41 GlnTrrIleValIleAspProLeuSerValIleTyrGluPheLeuThrAsnIleHisValPro 60
 Db 212 CAATGACCGCTCAAGCTTAAGACTGTCAAAATATGATTCACAGCTGATGTTACCTCC 271
 Qy 61 LysLeuGlyValMetLeuValGlyTrrGlyIleAsnAsnGlySerThrLeuThrGlyGly 80
 Db 272 AATTTAGGGGTTATAGCTTGTGGATGGGGTGAACAAATGCTTCAACCTTGACCGGTGGT 331
 Qy 81 ValIleAlaAsnArgGluGlyIleSerTrrAlaThrIleAspLeuIleGlnGlnAlaAsn 100
 Db 332 GTTATGCTTAACAGAAAGAAATTCATGCGCCACCAAGATAGGTGCAACAGCAAT 391
 Qy 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValIleGlySerPheGlnGlyGlu 120
 Db 392 TACTTTGGCTCTTACTCAGGCTTCTACTATTGAGTTGGGCTTTCAATGAGAAAG 451
 Qy 121 IleTyrAlaProPheLeuSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 Db 452 ATCTATGCTCCATTTAAAGCTCTCTTCCAAATGCTCAATCCAGATGACCTAGTGTGGA 511
 Qy 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
 Db 512 GAGTGGACATCAGCACAATGATTAAGCAATGCAATGCCAGGCTAAAGATTTAGAT 571
 Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
 Db 572 ATTGATCTCAAAAGAGAGTTGAGGCTCTACATGGAATCATATGATCCACCTCGGTATC 631

QY	181	YrAspPcRoAapPheIlaIaIaIaAsnGlnGluIaGAlaAaPnaValIleIySGly	200
Db	632	TATGACCCCTGATTATATGCTGCTAACAAGGCTGACGTGCACAAAGCTATCAAGCA	691
QY	201	ThrLySGInGInGInValGInGInIleIleIySaAPeIleIySaIaPheLySGInIaThr	220
Db	692	ACCAAGAAAGAAACAAATTGATCAATTCATTAAAGATATTAGGAGTTTAAAGAAAGAAC	751
QY	221	LySValAaPpLySValValIaIleuTrpThrIaAaSnThrGInIaGlyrSerAsnLeuVal	240
Db	752	AAAGTGGACAAAGTGCTGATGTGGCTGGACCTGCTAACACGTGAAGATACAGTAACTGTGCTT	811
QY	241	ValGlyLeuAaAaPpThrMeTcGInSnsnLeuIaIaIaValaSPaRGAaSnGInIaGIn	260
Db	812	GTTGACATTAAATGACACTRATGGAAAACTCTTGCTCTGTGGACAGAAATGAAGCTGA	871
QY	261	IleSerProSerThrLeuTrpAlaIleIaCySaVImeGInAaSnValProPheIleAsn	280
Db	872	ATATCTCTTCCACTTTGTATGACATTTGCTGACATCTTGAAAAAGTCCCTTTATTTAT	931
QY	281	GlySerProGInAaSnThrPheValProGlyLeuIleAaPLeuIaIleIaIaRGAaSnThr	300
Db	932	GGAAACCCCAAGAACACCTTGTGCCAAGCCCTCATTTGATTTGGCATGAAGAAGACACA	991
QY	301	LeuIleGlyGlyAaSPaPpPheLySereGlyGInThrLySmLySereValLeuValaSP	320
Db	992	TTGATTTGTGTGATGACCTTTAAAGGTGTGTCAACCAAAATGAATCATGCTGTGTGAT	1051
QY	321	PheLeuValGlyIaGlyIeIySPProThrSerIleValSerTyraSnHileuGlyAaSn	340
Db	1052	TTCCCTGTGTGGAGCTGTGATTAAGCCAAATCAATTGTGAGGTACAAACATTTGGGTAC	1111
QY	341	AaNaAPGlyMeTAsnLeuSerAlaProGInThrPheArgSerLySGInIleSerLySer	360
Db	1112	AATGATGAATGAATCATCTGTCCCTCAAACTTTCCGGTCAAAAGAAATCTCCAAAGAT	1171
QY	361	AaSnValAaSPaSPMeTValaSnSerIeAlaIeIeIyTrGInProGlyGInHilePro	380
Db	1172	AATGTGTGTGAAGACATGGTTCACCAAGATGCATCTTTATGAAGCTGTGAAGACACCT	1231
QY	381	AaPHisValIaValIleIySTyValProTyraValGlyAaSPeLySaRGAaMeTAsp	400
Db	1232	GACCATGTTGTTGTGATTAAGATATGTGCCATATGTGGAGACAGCAAGAGGCAATGAT	1291
QY	401	GInTyTrpThrSerGInIlePheMeTcGlyGlyLySerThrIleValLeuHileAsnThrCys	420
Db	1292	GAGTACACATCTGAGATTTTTCATGGGGGAAAGAACACACATTTGTTTCACAATCTTGT	1351
QY	421	GInAaSPeSerLeuLeuAlaIaProIleIleLeuAaPLeuValLeuLeuAlGInLeuSer	440
Db	1352	GAGGATTCACCTTTATGCTGCTCCCAATTAATTAATTAATTAATTAATTAATTAATTAAT	1411
QY	441	ThrArgIleGInPheLySaIaGInAaSnGInGlyLySPheHileSerPheHileProValaIa	460
Db	1412	ACCCGCATTTCAAGCTGAAGGTAAGAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTA	1471
QY	461	ThrIleLeuSerTyTrpLeuThrLySaIaProLeuValProProGlyTrpProValValaSn	480
Db	1472	ACTATCTCAGCTACCTTACCAAGGCTCCTGTGTAACACAGGTACACAGTGTGAT	1531
QY	481	AlaLeuSerLySGInArgAlaMeTcLeuGInAaSnIleMeTcGlyAlaCySaValGlyLeuAla	500
Db	1532	GCACTCTCAAAAGCAAGGCAATGCTTGAGAACATATTAAGAGGCTTGTGTGGACCTTCA	1591
QY	501	ProGInAaAaSnMeTleuGInTyTrpLyS	510
Db	1592	CCAGAGAACACATGATTTCTGSAATACAA	1621
RESULT 7			
LOCUS E27176 1950 bp DNA linear PAT 18-JUN-2001			
DEFINITION Novel INPS gene derived from micotiana.			
ACCESSION E27176			

VERSION	E27176.1	GI:13026394
KEYWORDS	JP 1997187879-A/1	
SOURCE	Nicotiana paniculata	
ORGANISM	Nicotiana paniculata	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Asteridophyta; Magnoliophyta; eudicotyledons; core eudicotyle; asterids; lamiales; Solanales; Solanaceae; Nicotiana.	
AUTHORS	Shigehiro, Y. and Toshiyuki, K.	
TITLE	Novel INPS gene derived from nicotiana	
JOURNAL	Patent : JP 1999187879-A 1 13-0UL-1999;	
COMMENT	JAPAN TOBACCO INC	
	OS Nicotiana paniculata	
	PN JP 1999187879-A/1	
	PD 13-0UL-1999	
	PF 26-DEC-1997 JP 1997359773	
	PR	
	PI SHIGEHIRO YAMADA, TOSHIYUKI KOMORI	
	PC C12N15/09//C12N5/10, (C12N15/09, C12R1:91), C12N15/00, C12N5/00,	
	PC (C12N15/00, C12R1:91)	
	CC Strandedness: Double;	
	CC Topology: Linear;	
	FM Key	
	FT CDS	
FEATURES	location/Qualifiers	
source	1..1950	
	/organism="Nicotiana paniculata"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:62141"	
ORIGIN		
Alignment Scores:		
Pred. No.:	2,75e-188	Length: 1950
Score:	2472.00	Matches: 470
Percent Similarity:	96.86%	Conservative: 24
Best Local Similarity:	92.16%	Mismatch: 16
Query Match:	93.92%	Indels: 0
DB:	6	Gaps: 0
US-10-718-952-2 (1-510) x E27176 (1-1950)		
QY	1 MetPheIleGIuAnbPhelSValGIuCySProaenValLySYrYThrGluThrGluIle	20
DB	92 ATGTTATTAGAGATTTTAAGGTGAGAGCCCAAGTTAAGTACACCGAAGTGAATT	151
QY	21 GluSerValTyrsnYrGluThrThrGluValuVLIsgLuAsnArgAsnGlyThrTyr	40
DB	152 CACTGTCTCATGATTATTAACAACCACTGAGTTGTCATGATGAGAAAAATGAGACATAT	211
QY	41 GluThrIleValLySProLySerValLySYrYrGluPhelSYThrAsnIleHisValPro	60
DB	212 CAATGACCGTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCC	272
QY	61 LysLeuGIyValMetLeuValGIyTYrGIyGIyAsnAsnGIySerThrLeuThrGIyGIy	80
DB	272 AAATTAGGGGTATGCTTGTGGATGGGGGTGAAACAATGTTCACCTTGACCGGTGT	333
QY	81 ValIleLeuAsnArgGIuGIyLysSerThrPalaThrLysAspLysIleGIuGlnAlaAsn	100
DB	332 GTTATTTGCTTACAGAGAAAGAAATTCATGAGGCGACCAAGAAATGAGTCAACAGCAAT	391
QY	101 TyrPheGIySerLeuThrGlnAlaSerAlaIleArgValGIySerPheGlnGIyGIuGIu	120
DB	392 TACTTTGGCTCTTACTAGGCTTCTACTATTCAGATTGGGTCTTTCAATGAGAAAGAG	451
QY	121 IleTYrAlaProPhelYSerLeuLeuProMetValaAsnProAspAspIleValPheGIy	140
DB	452 ATGTATGCTCATTTTAAAGCTCTTCCAAATGTGCATTCAGATGACGTAGTGTGTGGA	511
QY	141 GLyTYrAspIleSerMetLeuValaAspAlaMetAlaArgAlaLysValPheAsp	160
DB	512 GGAATGAGCAATCAAGCAATGAATTTAGAGATGCAATGAGCCAGGGCTTAAGGTATTTGAT	571

QY 116 AspLeuGlnIleuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 572 ATGATCTCAAAAGAGGTTGAGCCCTCATGGAATCATATGCTCCACCTCCCTGATC 631
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluIleValIleAsnValIleValGly 200
DB 632 TATGACCTGATTTCTGCTGTAACCAAGGCTCACTGCCCAACCAAGTGAACCAAGCA 691
QY 201 ThrIleGlnGlnGlnGlnGlnIleIleIleValAspIleValIlePheValIleThr 220
DB 692 ACCAAGAAAGAACAAATGATCAAAATCATTAAGGATTTAGGAGTTTAAAGAAAGAAC 751
QY 221 TyrValAspIleValIleValIleValIleValIleValIleValIleValIleVal 240
DB 752 AAAGTGAACAGGTGATGATTTGATGGAATGGAATGGAATGGAATGGAATGGAATG 811
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaValAspArgAsnGluIleVal 260
DB 812 GTTGAATTAATGACCTATGAAAACCTTTGCTTGTGAGCAAGAAATGAAGCTGAA 871
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValIleProPheIleAsn 280
DB 872 ATATCTCTTCCACTTGTATGCTATGCTGATGCTGAAATATGCTTTTATTAAT 931
QY 281 GlySerProGlnAsnThrPheValIleProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 932 GGAAGCCCCAGAACACCTTTGCTCCAGGCTCATTTGCTCCATCAAGAAAGAACCA 991
QY 301 LeuIleGlyIleAspAspPheIleValSerGlyIleThrIleMetLeuValIleValIleVal 320
DB 992 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1051
QY 321 PheLeuValIleValIleValIleValIleValIleValIleValIleValIleVal 340
DB 1052 TTCTCTTGTGAGCTGATTTAAGCCAAATGATGATGATGATGATGATGATGATG 1111
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerGlyIleSerLeuValIle 360
DB 1112 AATGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1171
QY 361 AsnValIleValIleValIleValIleValIleValIleValIleValIleValIle 380
DB 1172 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1231
QY 381 AspPheValIleValIleValIleValIleValIleValIleValIleValIleVal 400
DB 1232 GACCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1291
QY 401 GlyTyrThrSerGlnIlePheMetGlyIleValSerThrIleValIleValIleValIle 420
DB 1292 GAGTACACATCTGAGATTTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1351
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleLeuValIleLeuVal 440
DB 1352 GAGGATTCATTTTATGCTCTCCAAATTAATGATTTGCTCTCTCTCTCTCTCTCT 1411
QY 441 ThrArgIleGluPheValIleValIleValIleValIleValIleValIleValIleVal 460
DB 1412 ACCCGCATTCACCTCAAGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1471
QY 461 ThrIleLeuSerTyrIleThrIleValIleValIleValIleValIleValIleValIle 480
DB 1472 ACTATCTCAAGTACTTTCACCAAGGCTCTCTGATGACCAAGGATCAACAGGAGGAG 1531
QY 481 AlaLeuSerIleGlnArgIleMetLeuGluAsnIleMetArgAlaCysValIleValIle 500
DB 1532 GCACTCTCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1591
QY 501 ProGluAsnAsnMetIleLeuGlyTyrIleVal 510
DB 1592 CCAAGAAACAAATGATTTCTGAAATTAACAA 1621

RESULT 8

BT013505
LOCUS BT013505 1959 bp mRNA linear PLN 11-MAY-2004
DEFINITION Lycopersicon esculentum clone 132193F, mRNA sequence.
ACCESSION BT013505
VERSION BT013505.1 GI:47104920
KEYWORDS FLI CDNA.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons; Asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 1959)
AUTHORS Kirkness, E.F., Wang, W. and Vazeille, A.
TITLES Direct Submission
JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
1..1959
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132193F"
/issue_type="Final"
/note="TXBFS1"
ORIGIN
Alignment Scores:
Pred. No.: 4,81e-188 Length: 1959
Score: 2469.00 Matches: 472
Percent Similarity: 96.27% Conservative: 19
Best Local Similarity: 92.55% Mismatches: 19
Query Match: 93.81% Gaps: 0
DB: 8
US-10-718-952-2 (1-510) x BT013505 (1-1959)
QY 1 MetPheIleGluAspPheValIleGlyCysProAsnValIleTyrGlnThrGlnIle 20
DB 77 ATGTTATGAAATTTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 136
QY 21 GlnSerValTyrAsnTyrGlnThrThrGlnLeuValIleGluAsnArgAsnGlyTyrTyr 40
DB 137 CATTCGTGTATGATTAATTAACCAAGAGCTGTTTCATGAGAGAGAGAAATGAACTTAT 196
QY 41 GlnTPIleValIleAspProIleValIleValIleValIleValIleValIleValIle 60
DB 197 CAATGATGTTTAAGCTTAACATGTCAAATATGAATTTAAATGATCCCATGCGCA 256
QY 61 LysLeuGlyValMetLeuValIleTyrProGlyIleValAsnGlySerThrLeuThrGlyIle 80
DB 257 AAATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 316
QY 81 ValIleAlaAsnArgGlnGlyIleSerTyrAlaThrIleAspIleIleGlnGlnAlaAsn 100
DB 317 GTTATTTGCAATCGAAGAGATTTTCATGCGACAGAAAGAAAGAAAGTGCACAGCCAT 376
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValIleGlySerPheGlnGlyGln 120
DB 377 TATTTTGGGTCTCTTAATCAAGGATCAACCATTCAGATGGGTCTTTCAATGGCGAAG 436
QY 121 IleTyrAlaProPheIleSerLeuLeuProMetValAsnProAspAspIleValIlePheGly 140
DB 437 ATCTATGACCCCTTCAAAAGCTCTCTCCATGCTCAACCCAGACGATGATGATTTTGA 496
QY 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValIlePheAsp 160
DB 497 GGAATGACATTAAGCAACATGATTTTGGCAATGCTATGCTAGAGGCTTAAGTTTCGA 556
QY 161 IleAspLeuGlnIleGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 557 GTTGATCTGCAAAACAGCTGAGGCTTCAACAGGATCATGATGATGATGATGATGATG 616
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

```
Db 617 TATGACCCCTGACTTCATTCGGCTAAACAAGACGCGCAACAACCTGTCAAGA 676
Qy 201 ThrTyrGlnGluValGlnGlnIleIleLysAspIleLysAlaPheIleGluIleThr 220
Db 677 ACCAAGAAAGAACAAAGTTGAACAATGTTAAAGATTAATTAAGAGATTCAAGAGAAAC 736
Qy 221 LysValAspLysValValIleuThrThrAlaAsnThrGluuArgTyrSerAsnLeuVal 240
Db 727 AAGTAGACAAAGATAGTGGTTCATAGGACTGGCAACCGAAGGATAGATATGGT 796
Qy 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaValAspArgAsnGluIleGlu 260
Db 797 GTTGCGCTTAATGACACCAATGAAAACTTTAGCTGCGTGGATAGAAATAGCGCTGA 856
Qy 261 ILeSerProSerThrLeuThrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 857 AATATCTCTTCTACATTCGATGCTATGCTGTATATGAAAAGTGCCTTTCATCAAC 916
Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 917 GGAAGCCCTCAAAACAACCTTTGTTCCAGGCTCTTATGATTTGGCCATAAAGAAACA 976
Qy 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 977 TTAATTTGGTGTGATGACCTTTAAGAGTGGTCAAAACCAAGATGAAAGTCCGTGTGTTGAT 1036
Qy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleuGluAsn 340
Db 1037 TTCTCTGTGTGAGACCGGATATTAAGCCACGTCATATGATAGCTACATACATCTGGGTAC 1096
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db 1097 AATGATGGAATGATATCTTCTGCTCTCAACCTTCGGTCTTAAGAGATCTTCAAAAGT 1156
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db 1157 AATGTTGTGATGACATGTTGCTAGCAACGCCATCTCTATGAAATCTGGCGACACCT 1216
Qy 381 AspHisValValIleLeuThrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1217 GACCATGTTGTGATGATCAAGATGTTCCATATGTTGAGACGACGAGGCGCAATGAT 1276
Qy 401 GluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db 1277 GAGTACATCTCAGAGATTTTTCATGGCGGAAAGACACTATAGTTATGCAACAACCTTGT 1336
Qy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1337 GAGGACTCTCTTTGGCACCTCCAAATTAATCTTGATTTGCTCTCTCCGTAACCTCAC 1396
Qy 441 ThrArgGluGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db 1397 ACTCGCATTCAGCTCAAGAGCTGAAGGAGGAGGAGATTCACCTCTTCCACCTGTGGCT 1456
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1457 ACTATTCTCAGCTACTCAACAAGGCTCCCTGGTATCCACCAAGTACTCAAGTGGTGAAT 1516
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1517 GCCCTTTCAAGCAGAGGCGCAATGCTTGAGAAATATATGAGAGCTGTGTGTGTGGCA 1576
Qy 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1577 CCAGAGAACACATGATTTTGGAAATCAAA 1606

RESULT 9 AF284065 1845 bp mRNA linear PIN 30-JUL-2003
LOCUS Sesamum indicum myo-inositol 1-phosphate synthase mRNA, complete
DEFINITION Cds.
ACCESSION AF284065
```

```
VERSION AF284065.1 GI:9858815
KEYWORDS Sesamum indicum (sesame)
SOURCE Sesamum indicum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Pedaliales; Pedaliaceae; Sesamum.
REFERENCE 1 (bases 1 to 1845)
AUTHORS Chun,J.A., Jin,U.H., Lee,J.W., Yi,Y.B., Hyung,N.I., Kang,M.H.,
Pyee,J.H., Suh,M.C., Kang,C.W., Seo,H.Y., Lee,S.W. and Chung,C.H.
Isolation and characterization of a myo-inositol 1-phosphate
synthase cDNA from developing sesame (Sesamum indicum L.) seeds:
functional and differential expression, and salt-induced
transcription during germination
JOURNAL Planta 216 (5), 874-880 (2003)
MEDLINE 22511750
PUBMED 12624775
REFERENCE 2 (bases 1 to 1845)
AUTHORS Jin,U.-H. and Chung,C.-H.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2000) Division of Biotechnology, Faculty of Life
Science and Resources, Dong-A University, Ha-Dan-Dong, Sa-Gu-Gu,,
Pusan City, Pusan 604-714, South Korea
FEATURES
source location/Qualifiers
1..1845
/organism="Sesamum indicum"
/mol_type="mRNA"
/db_xref="taxon:4182"
/tissue_type="seeds"
/note="authority: Sesamum indicum L."
51..1583
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAG01148.1"
/db_xref="GI:9858816"
/translation="MFIRSFVSEPNVYKGEIHSYVNYETTELVHSENGTYOMIV
KPKTVKFKTDIVPKLGLVMVGKNGSTLTGVIANRSGISMAKDKYQQANF
GSLTQASSIRVGSFNGEBEIVAFKSLILPMVNDVDVFGWMISSNNLADAMGRVALD
IDLOKOLRPMHWPPLGIVDPDFIANQGSRAANNVTKGRKEQVOQIIDMRPKE
ONKVDKVVLVLTANTERSNVVGLNDPESISVENEAISPTLYAIAQVENV
PRINSPOINTFVPLIDAIORNSLIGDDPSGQTKRKSVALDELVAGIKPTISV
YNHLGNNGMNLASQTRSKSISNVVDVMVNSGILYRGHRDHIIVIKVYPV
GDSKQAMDEYSEIFMGKSTIVLHNTCEDSLIAPILDLVLABELSTRIDLRKEG
GKFSFHFVAITLSTYLRAPLVPPTPVNALSKQRAWLENILRACVGLAPENNMLE
YK"
ORIGIN
Alignment Scores:
Pred. No.: 8,53e-187 Length: 1845
Score: 2453.00 Matches: 470
Percent Similarity: 96.08% Conservative: 20
Best Local Similarity: 92.16% Mismatches: 20
Query Match: 93.20% Indels: 0
Gaps: 0
DB:
US-10-718-952-2 (1-510) x AF284065 (1-1845)
Qy 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
Db 51 ATGTTCAATCGAAGAGCTTCAAGGTTAGAGAGCCCAATGTCAGTACCAAGGATAGATT 110
Qy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 111 CACTCGTCTACCAACTAGAGACTACAGAGCTTGTTCATGAGAGCAGAAAACGGACACTAC 170
Qy 41 GlnTPDIIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db 171 CAATGATTTGTCAAACCCAGAGCTGTCAAAATACGATTCAAACATGATATGTATGCC 230
Qy 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 231 AATATTAGGGGATCATCTGTTGATGGAGGAGGAAACATGTGTTCAACTCTCACTGGCGGT 290
```


OY	81	ValIleAlaAsnArgGlnGlyIleSerThrAlaThrLysAspLysIleGlnGlnAlaAsn	100
Db	231	GTCATTGCCAACAGGAAAGGATTTCAATGGCCAAAGAGATTAAGTGTCAACAGCCCAAT	350
OY	101	TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlnGlu	120
Db	351	TATTTTGGGTCACCTGACACAGGCAATCTTCATCAAGATTGGTCTCTTCAATGAGAAAG	410
OY	121	IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValIleGly	140
Db	411	ATCTATGCCCTTCTCAAAAGCTTCTTCCCATGCTCAACCCAGATGAGTGTGTTGGG	470
OY	141	GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp	160
Db	471	GGATGGGACATCAGCAATATGTAATTGGCTGTATGGCTATGGGCAAGGCGCAAGGTGTTAAT	530
OY	161	IleAspLeuGlnLysGlnLeuArgProTyrMetGlnSerMetLeuProLeuProGlyIle	180
Db	531	ATTGATCTCCAGAAACAGCTCAAGGCTCTACATGAGGACATGTGTCCACTCCCTGGAAATC	590
OY	181	TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly	200
Db	591	TACGATCTCGATTTTCATATGACGCGCAATCAGGGATCACTGCGCCAAACAAGTATTAAGA	650
OY	201	ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr	220
Db	651	ACCAAGAAAGAAACAAGTTTCACAGATCATCAAAAGCATGAGGAGACTTTAAGAGGCAAAAT	710
OY	221	LysValaAspLysValaValLeuThrPheAlaAsnThrGlnLysTyrSerAsnLeuVal	240
Db	711	AAAGTGGACAAAGTCTGATGCTGTGTGACACGCCAACACAGAAAGTATCAGTATGTTGTC	770
OY	241	ValGlyLeuAsnAspThrMetGlnAsnLeuLeuAlaAlaValAspArgAsnGluAlaGlu	260
Db	771	GTTGGGCTAAACGATACAGCCGAAAGCTGATGCGCTCGGTAGAGAGAAATGAGGCCGAG	830
OY	261	IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn	280
Db	831	ATATCTCTTCAACCTGTATGCCCATTGCTGTGTTTGAAGAAATGTTCCGTCTTATAT	890
OY	281	GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr	300
Db	891	GGCAGCCCAACAAATATCTTTGTTTCCAGGGTTGATGATTGGCGATTCAGAGGAACGT	950
OY	301	LeuIleGlyGlyAspAspPheLysSerGlyIleThrLysMetLysSerValLeuValAsp	320
Db	951	TTGATGGTGGAGATGATTTCAAGAGCGGTCAACCAAGATGAAATCATGCTGATGAT	1010
OY	321	PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleGlnLysAsn	340
Db	1011	TTCTCTTGTGAGCTGGTATCAAGCCAAACGTGATTTGAGCTTAATCATCTGGGAAC	1070
OY	341	AsnAspGlyMetAsnLeuSerValProGlnThrPheArgSerLysGlnIleSerLysSer	360
Db	1071	AACGACGGGAATCTATCAAGACCCCAACTTTCGCGTCCAAAGAAATTTCTTAAAGC	1130
OY	361	AsnValValaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnIlePro	380
Db	1131	AATGTCGTAGACGACATGTGGTGGCCAGCAAGCGCACTTTATGAGCCTGGTGAACACCC	1190
OY	381	AspIleValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp	400
Db	1191	GACCATATATTTGTCATCAAGTATGTAACCTATGTGGGGGACAGCAAGAGGCAATGCAC	1250
OY	401	GluTyrThrThrSerGlnIlePheMetGlyGlyLysSerThrIleValLeuAsnThrCys	420
Db	1251	GAGTACACTTTCGAGATATTCATGGGAGGAGAAAGTACCAATGTTCTGCACACACCGGT	1310
OY	421	GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer	440
Db	1311	GAGGACTCTCTTGGCTGGCTCCATCATCTCTGATTTTGGTCTCTCTTAAGCTGAATCAGC	1370
OY	441	ThrArgIleGluPheLysAlaGluAsnGlnGluLysPheIleSerPheIleProValAla	460

Df		1371	ACTGCTATTACAGCTTCAAAGCCGAGGAGGAGGGCAAAATTCATTGCTTCCACCCTGTGCA	14330
Oy		461	ThrlleuseSerryleuthrlygalaProleuvaIProProgllythrProvalValasn	480
Df		1431	ACCATCCTCAGACTACCTAACCAAAAGCCCCTTTTACTCTCCAGGCACAACCGTGCTAAC	1490
Oy		481	AlaleuserlyvglnaAgalAmetleglunasnllmetcatalaCyvalaglyleuala	500
Df		1491	GCACCTGCTTAAGCAGCGTCCACTCTCGAGAACACTTGAAGGCGCTGTGTGATGGCT	1550
Oy		501	ProglunabnmetlleuglnTyrrlys	510
Df		1551	CCAGAAAACATATGATTTTGGAAATACAG	1580
RESULT 10				
LOCUS		BT013759	1954 bp	mRNA linear PLN 11-MAY-2004
DEFINITION		Lycopersicon esculentum clone 132642F, mRNA sequence.		
ACCESSION		BT013759		
VERSION		BT013759.1	GI:47105174	
KEYWORDS		FLI CDNA.		
SOURCE		Lycopersicon esculentum (tomato)		
ORGANISM		Lycopersicon esculentum		
REFERENCE		Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.		
AUTHORS		1 (bases 1 to 1954)		
TITLE		Kirkness, E.F., Wang, W. and Vazellile, A.		
JOURNAL		Direct Submission		
FEATURES		Submitted (11-MAY-2004) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA		
source		Location/Qualifiers		
		1..1954		
		/organism="Lycopersicon esculentum"		
		/mol_type="mRNA"		
		/db_xref="taxon:4081"		
		/clone="132642F"		
		/tissue_type="fruit"		
		/note="TZXBV41"		
ORIGIN				
Alignment Scores:				
Pred. No.:		1,21e-185	Length:	1954
Score:		2439.00	Matches:	463
Percent Similarity:		95.88%	Conservative:	26
Best local Similarity:		90.78%	Mismatches:	21
Query Match:		92.67%	Indels:	0
DB:		8	Gaps:	0
		US-10-718-952-2 (1-510) x BT013759 (1-1954)		
Oy		1	MechehellEgluanenPhelysvaIGluCyPProasnVallystrYThrIngluthrcIUlle	20
Df		107	ATGTTTATTTAGAACTTTAAGTTGAGAGGCCCAATGTTAAGTACACAGAGAGTAGATT	166
Oy		21	GlnserValTYraenTYrgIutnrThcGluevahiGlnunaArgasngLYthTyx	40
Df		167	CACCTCTGTrnATnTTCACAAACACGAGTGGTTTCATAGAGAAAAAATGGACATT	226
Oy		41	GlnTrpillevallysProlysSerVallystrYrgIubhelysrThraenllehisvalpro	60
Df		227	CAATGACCTGTAAGCCCAAACACTGCAATATGAAATTCAAAACCTGATGTTCAATTCCA	286
Oy		61	LytleuglyValmetleuvaiGlyTYrrnglyGlyAsnasnglyserThrleuthrlyGly	80
Df		287	AAATTTAGAGATTTATGCTTGTGATGGGGTGGAAACATGTTCAAGCTTGACCGAGGT	346
Oy		81	ValllleAlaasnArggluglylieserTPAlarThrysAsplysllleglnglnAlaasn	100
Df		347	GTTATTCTTAATAGAGAGAAATTTCAATGGGCTTACCAAGATAAAGTGCAGCAAGCCAAAT	406
Oy		101	TyPheclYserleuthrlnalAservalIleArgValglyserPheginglyGlu	120

```

Db      407 TACTTGGTTCTTACACAGGCTCTACTATTGAGTTGGGCTTTTCAATGAGAGAG 466
Qy      121 IleyrAlaProPhelyserleuLeuPrometValAsnProAspAlaIleValPheGly 140
Db      467 ATCTATGCTCCCTTTAAAGACATCTCTCCATGGTCAATCCAGATGATGATTTTGGAA 526
Qy      141 GATTPAPAPILeserAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
Db      527 GGATGGGACATACGACACATGATTTGGCAGATGCCATGGCAAGACTTAAAGTTTGGAC 586
Qy      161 ILeasPLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db      587 ATTGATCTCTCAATGCAATGCACTGAGGCGGTACATGAAATCCATGATGCCACTGCTGATAC 646
Qy      181 TYRAspProAspPheIleAlaIleAsnGlnGluArgAlaAspAsnValIleLysGly 200
Db      647 TATGACCTGATTTCACTGCTCCAAACCAAGGCTCAGTGTACAAATGTCACAAAGGA 706
Qy      201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db      707 ACCAAGAAAGAACAAATTATCATCAATTAATTAAGATATTAGGAGCTTTAAGAGAACAAAC 766
Qy      221 LysValAspLysValValIleLysPheAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      767 AAAGTGAAGACAGGTGGTGGTCTCTGTCGACCTGCCAACACATGAAAGATACAGCTGGCT 826
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGlnAlaGlu 260
Db      827 GTTGGCTTTAATGATACCAATGAGAAACCTTTGCTGTGATGATGAAATAGAGGTGAA 886
Qy      261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      887 AATCTCTCTTCCACCTGTGATCTGATGCTTGTGATCTTGAAATGATACCTTCATCAAT 946
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      947 GGAAAGCCCAAAATACTTTTGTCCAGGCTCATTTGATTTGGCCATCAAGAGAACACT 1006
Qy      301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      1007 TTGATTTGGTGGTGTACTTTTAAAGTGTGTCAAAACCAATGAAATGACGTGGTGGTAA 1066
Qy      321 PheLeuValGlyValAglyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db      1067 TTCCTTTGGTGAAGCTGTATTAAGCCAAATCATGATGATGATCAACCATTTGGTAAAC 1126
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db      1127 AATGATGAAATGAAATCTGTCTGCCCCCTCAAACTTTCCGTTCAAAAGAGATCTCGAAAAAC 1186
Qy      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnHisPro 380
Db      1187 AACGTTGTGATGACATGCTTTCTAGTAAACGCATCTTTATATAGCTTGAAGAGCATCTCT 1246
Qy      381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1247 GACCATGTTGTTGTATTAAGTACGTAACCATATGTCGAGATGACAAAGCGCAATGAT 1306
Qy      401 GlnTyrThrSerGlnIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db      1307 GAGTACACATCAGAGATTTTCACTGAGTGAAGGAACACATGTATTTGCAACAATACCTGT 1366
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1367 GAGGATTTCTTTTGGCTCTCCAAATTAATCTTGATTTAGTGTCTCTGCTGAACTCACT 1426
Qy      441 ThrArgIleGluPheLysAlaGluAsnGlnGlyLysPheHisSerPheHisProValAla 460
Db      1427 ACCCGCATTCAGCTCAAGCTGAAGAGAGGGTAAATTCCACTCTTCATCCCGTGGCG 1486
Qy      461 ThrIleLeuSerTyrIleuThrLysAlaProLeuValProProGlyThrProValValAsn 480

```

```

Db      1487 ACCATCTCAGTTATCTTACCAAGGCTCTTGGTACCAACAGGTACACCACTGCTAAT 1546
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1547 GCCCTTTCAACACAGAGGCAATGCTCTGTAACATATTAAGAGGCTTGTCTTGAAGTGGCA 1606
Qy      501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db      1607 CCAGAGAACACATGATCTTGGATATCAAA 1636

RESULT 11
AF120146      1538 bp      mRNA      linear      PLN 23-APR-1999
LOCUS
DEFINITION
Triticum aestivum cultivar Fielder myo-inositol 1-phosphate
synthase (MIPS) mRNA, complete cds.
ACCESSION
AF120146
VERSION
AF120146.1 GI:4589061
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 1538)
Hussein, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
CDNA clone for myo-inositol 1-phosphate synthase from wheat
Unpublished
JOURNAL
2 (bases 1 to 1538)
Hussein, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
Direct Submission
Submitted (14-JAN-1999) National Research Council Canada, Plant
Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N
0W9, Canada

FEATURES
source
1..1538
location/Qualifiers
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Fielder"
/db_xref="taxon:4565"
1..1538
/gene="MIPS"
1..1533
/gene="MIPS"
/EC_number="5.5.1.4"
/function="converts glucose 6-phosphate into myo-inositol
1-phosphate"
/note="cytosolic enzyme"
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAD26330.1"
/db_xref="GI:4589062"
/translacion="MIPSFKVSPNVKYTEHSHSVYDYETTEVVENRNGTYQWV
KPKTYKPKFDTRPKGLVMDVGKNGSLTGTGVANKEGISMATDKYQOANVF
GSLTVAASSIRGSIVGEBEIVAFKSLPMVNPDDVDFGMDISDNMLADNARAYLD
IDLQKLRPEHWPVLPGLIDPDIANQGSRAVSVKTKKEOVDOITIKMREFE
KMKDKLVLTMTANERYSDDVVGLNDMENTLASEVKNESISPSITVALA CVLEGI
PINGSPPNTFVPGILELISKNCLIGDDFKSGTQKSVLVDPLVNGIKPSTIVS
VNHLDNNGMNLASQTRSKSISKSNVDDVAVANGLIEBGEHPDHVVVXYVPYV
GDSKRAMDYETSEIFMGKNTVMNTEDESLIAPILIDLVILAEISRIOFKAGE
GKFSHFHFAITLSTLTATPLVPACTPVVNALSKORAMLENTLRACVGLAPENNMLE
YK"

ORIGIN
Alignment Scores:
Pred. No.:      5,2e-182      Length:      1538
Score:          2392.00      Matches:      453
Percent Similarity: 95.88%      Conservative: 36
Best Local Similarity: 88.82%      Mismatches: 21
Query Match:     90.88%      Indels:      0
DB:              8          Gaps:      0

US-10-718-952-2 (1-510) x AF120146 (1-1538)

```

Qy	1	MetPheIleGIuAenPheIyValGIuCySProAenValIySyrThGluThGluThIle	20
Db	1	ATGTCATCGAGCGCTTCAAGATTGAGAGCCCAAACTGTAAGTACACAGAGAACGAGATC	60
Qy	21	GIuSerValIyYAsnTyGluThrThrGluLeuValIhIeGIuAsnAxAbsnIyThrTyx	40
Db	61	CATTCTGTGTACGATTACGAAACCGAGAGCTGCTCATGAGAACCTGTACCGCACTTAC	120
Qy	41	GIuTrpIleValIySProlYsSerValIySyrTyGluPheIyThrAsnIleHIVaIlePro	60
Db	121	CAATGGCGTCACAGCCCAAAACCTGCAAAATACGATTCAAAACAGACACTGCGTCCC	180
Qy	61	LysIleuGIyValIyMetLeuValGIyTrpGIyGIyAaAsnGIySerThIleuThrGIyGIy	80
Db	181	AAGCTAGGCGTTATGCTGTGGATGGGAGGAACAAATGATCAACTTAACTCCCTGGA	240
Qy	81	ValIleAlaAsnAxAbsnGIuGIyIleSerTrpAlaThrIyAspIyAlIeGIuAlaIaAsn	100
Db	241	GTGATCGCCAAACAAAGAGGATCTCGTGGGACCTTAAGACAAAGTGCACACACAAAC	300
Qy	101	TyrPheGIySerLeuThrGluAlaSerAlaIleArgValGIySerPheGIuGIyGIuGIu	120
Db	301	TACTTGGCTCCCTCATTCAAGCTTCTTCGATCCGGGTGGATCTTACAAATGTGAAGG	360
Qy	121	IleTyTrAlaProPheIySerIleuLeuProMetValIaAsnProAspIleValIyPheGIy	140
Db	361	ATCTACGCTCTTTCACAGAGCTCTCCCTCCATGGGAACCTCGACGATGTGTGTGTGT	420
Qy	141	GIyTrpAspIleSerAsnMetAsnIleuAlaAspAlaMetAlaArgAlaIyValIyPheAsp	160
Db	421	GATTGGACATAGCGCACATGACTTACGACGCTATGCTAGACGACAGAGTCTTTGATC	480
Qy	161	IleAspLeuGIuIyGIuLeuArgProTyrmecGIuSerMetIleuProLeuProGIyIle	180
Db	481	ATTGATCTCCAGAAACAGACTGAGCGCTGATGAGACATGAGTCCACTCCCTGTATC	540
Qy	181	TyrAspProAspPheIleAlaIaAsnGIuGIuAlaIaAsnValIleIyGIy	200
Db	541	TACGACCTGTACTCATGCTGCGCATCAAGGGTCACCTGCTAAATAGGTGATCAAAAGT	600
Qy	201	ThrIySGInGIuGIuValGIuGIuIleIleYsAspIleIyValIyPheIyGIuAlaThr	220
Db	601	ACCAAGAAAGACAGATGACACCGATATCAAAAGCATAGAGGAGCTTAAAGAGAAAGAC	660
Qy	221	LysValIyAspIyValIaIleuTrpThrAlaAsnThrGIuArgTySerAsnIleuVal	240
Db	661	AAGGTGATTAAGCTGCTGTCTGTGACACGACACACAGACCTTACAGCATGTGCTC	720
Qy	241	ValGIyLeuAsnAspThrMetGIuAsnIleuAlaIaValIyAspArgAsnGIuAlaGIu	260
Db	721	GTAGGGCTTAACGACACATGAGACATCTATTGCTGCTGTGGAAGAACAGACTGAG	780
Qy	261	IleSerProAspThrIleuTyTrAlaIleAlaCyAsValIyMetGIuAsnValIyProPheIleAsn	280
Db	781	ATCTGCTCTTACGCTATATGATGCGATGCTGTGTCTTGAAGCATTCCTTCAATAT	840
Qy	281	GIySerProGIuAsnThrPheValProGIyLeuIleAspIleuAlaIleAlaArgAsnThr	300
Db	841	GGAGTCCGACAAACACTTGTCTTCTGACTCATCGAGCTGGCTATATCGAAGAACTGT	900
Qy	301	LeuIleGIyGIyAspAspPheIySersGIyGIuThrIySmetIySersValIleuValAsp	320
Db	901	TTGATCGGTGGATGATCTTCAAGAGTGTGACATTAAGATGATGTTTGGTGTAT	960
Qy	321	PheIleuValGIyAlaGIyIleIyProThrSerIleValIySerTyTrAsnHileuGIyAsn	340
Db	961	TTCTCTTGGAGCGGACATCAAGCTTACATTGAGCTATTAATCAATTGGGAAAC	1020
Qy	341	AsnAspGIyMetAsnIleuSerAlaProGIuThrPheArgSerIyGIuIleSerIySers	360
Db	1021	AATAGCGGATGAACTCTTCAAGCGCTCAAGCTTTTAAATCTTAAGAAATCTCAAGAGCC	1080
Qy	361	AsnValIyAspAspMetValAsnSerAsnAlaIleIleuTyGIuProGIyGIuHilePro	380

[illegible]

Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N 0M9, Canada
Location/Qualifiers
1. .1538
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Biggar"
/db_xref="taxon:4565"
1. .1538
/gene="MIPS"
1. .1533
/gene="MIPS"
/EC_number="5.5.1.4"
/function="converts glucose 6-phosphate into myo-inositol 1-phosphate"
/note="cytosolic enzyme"
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAD26332.1"
/db_xref="GI:4589066"
/translation="MTESFKEBPVVKTTENEHSVDYETTEVHNKRGTYQNV
KRTVAYDKTDTRVPLKGLVGLVWGNGNSTLTAGVANKGISWATDXYQANVF
GSLTQASIRVGSYNGEITAPFKSLPMVNPDDVFGMDISDMNLADAMARVLD
IDLOKLRPYMEHNVPLPGIYDPDFIAQNGSRANSVIGTKKEVDIOIIXMBREKE
KKKVDLVVLTANTERYSDVVVGLNDTMENTNLASVEKNESISPSLTVAIAVCYEG
PINGSPOPTFVPGELIALSKKCLIGGDFKSGQTKMSVLDPLVGCIPTSTVS
YHKLNDMDNLSDAPQTRSKLSKSNVDDVAVSGLLPEGEHDDHVVILKYEV
GSKAMDYTSIEIPWGKNTIVMNTCDSLLAPRIIDLVLMLSTRIOFKEGE
GKFHSFVPATILSYLTAKPLVPAGTPVNVNLSKORAMENILRACVGLAPENMLLE
YK"

ORIGIN

Alignment Scores:
Pred. No.: 5.2e-182 Length: 1538
Score: 2392.00 Matches: 453
Percent Similarity: 95.88 Conservative: 36
Best Local Similarity: 86.82 Mismatches: 21
Query Match: 90.88 Indels: 0
Dn: 8 Gaps: 0

US-10-718-952-2 (1-510) x AF120148 (1-1538)

Qy 1 MetPheileGluAsnPhelValGluCyProAsnValIleThrGluThrGluile 20
Db 1 ATGTTTCATGAGAGCTTCAAAGTTGAGAGCCCAAGCTAAAGTACCAAGAA CGAGATC 60
Qy 21 GlnSerValIleAsnIleThrGluThrGluLeuValHleGluAsnArgAsnGlyThrTy 40
Db 61 CATTCGTGTTACGATTACGAACCAAGCGTGTCTCATGAGAACCTTAACCGCATTC 120
Qy 41 GlnTPHleValIleProIleSerValIleTyGluPhelGlyThrAsnIleHleValPro 60
Db 121 CATGAGTGTGTCAGAGCCCAAAACCGTCAATACGATTCAAAAACAGACATCGCGTCCC 180
Qy 61 LysLeuGluValMetLeuValGluTyProGluValAsnAsnGlySerThrLeuThrGly 80
Db 181 AAGCTAGGCGTTATGCTGTGATGAGGAGAAACAATGATCAACCTTAACCGCTGGA 240
Qy 81 ValIleAlaAsnArgGluGluIleSerTPAlaThrIleAspIleIleGlnIleAlaAsn 100
Db 241 GATGATGCCAACAAAGAGATCTGTGGCGCATTAAGACAAAGTGCACAGCAAGAAC 300
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTTTGGCTCCCTCACTCAAGCTTCTTGATCCGGGTGGATCTTAACAATGGTGAAG 360
Qy 121 IleTyrAlaProPheIleSerLeuLeuProMetValAsnProAspIleValIlePheGly 140
Db 361 ATCTACGCTCTTTCACAGAGCTCTCCCTCATGATGTAACCTGACCATCTTGTTGGT 420
Qy 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValIlePheAsp 160
Db 421 GGTGGGACATTAAGCAATGAATCTTAAGCAAGCGCTATGGCTAGCAAGAGATCTTGGAC 480

Qy 161 IleAspLeuGlnIleGlnIleuArgProTyMetGluSerMetLeuProIleProGlyIle 180
Db 481 ATTGATCTTCACAAACAGCTCAGGCCGTATCAGAGCACATATGCTCCCTGGATATC 540
Qy 181 TyrAspProAspPheIleAlaIleAsnGlnGluIleArgAlaAsnAsnValIleTyGly 200
Db 541 TACGACCTGATCTTCACTCCCTCGAATCAAGGTCACGTCCTTAATAGCTGATCAAGGT 600
Qy 201 ThrIleGlnGluIleValGlnIleIleIleTyAspIleLysAlaPheIleGluIleThr 220
Db 601 ACCAAGAAAGAACCAAGTCACACAGATCATCAAAAGCATAGAGGAGATTTAAGAGAAAGAAC 660
Qy 221 LysValAspIleValValIleuThrPThrAlaAsnThrGluIleTySerIleAsnVal 240
Db 661 AAGGTGAGTAACTGGTGTCTTGAGACAGCAACAGACCGTTACGCGATGTGGTGC 720
Qy 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaIleValAspArgAsnGluIleGlu 260
Db 721 GATGGCTTTACGACACGATGAGAAATTAATGGCTTCTTGAGAGAAACGATCTGAG 780
Qy 261 IleSerProSerThrLeuTyAlaIleAlaCyValMetGluAsnValProPheIleAsn 280
Db 781 ATCTGGCTTCACTACGCTATATCGATGCTGTGTTGTTGTAAGGACATTCCTTCATTAAT 840
Qy 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGGAGTCCCAAGAACACGTTTCTCTGACATTCAGCTGCTCATATATCAAGAACTGT 900
Qy 301 LeuIleGlyIleAspAspPheIleSerGlyIleThrIleMetLysSerValLeuValAsp 320
Db 901 TTGATCGGGGAGATGACTTCAAGAGTGTCAAGATAGATAGATAGCTGTTTGGTGAAT 960
Qy 321 PheLeuValGlyAlaGlyIleIleTyProThrSerIleValSerTyrAsnIleIleGluIleAsn 340
Db 961 TTCCTTGTGGAGCGTGGATCAAGGCTACTTGATTTGAGCTGTGATATATCATTTGGGAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGluIleSerLysSer 360
Db 1021 AATGACGGCATAACTCTCAAGCGCTCAGACTTTTATGATTCAGAGAGTCTCAAGAAC 1080
Qy 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyGluProGlyGluIleAspPro 380
Db 1081 AATGTTGTTGACAGATGCTGCTAGTACGCTATCTTTGAGCTGTGTAAACATCTCT 1140
Qy 381 AspHleValIleValIleIleTyValProTyValIleGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCATGTTGTGTCATCAAGATGTCTCTTAATGTTGGGATAGTAAAGAACCATTTGAC 1200
Qy 401 GluTyThrSerGluIlePheMetGlyIleLysSerThrIleValIleuHleAsnThrCy 420
Db 1201 GAGTACACCTCCGAGATATTTCTAGGAGGAAGAAATACATTTGTTATGATCAACA 1260
Qy 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleLeuAlaGluLeuSer 440
Db 1261 GAGGATCTCTTTTGGCTCTCAATCAATCTTGATCTTGTCTTCTCGCTAGCTCAGC 1320
Qy 441 ThrArgIleGluPheLysAlaGluAsnGluIleLysPheHleSerPheHleProValAla 460
Db 1321 ACCAGAAATTCATTAAGAGCTAAGAGGAGGAGGAGTTTCATCTTTCCACCAAGTACGA 1380
Qy 461 ThrIleLeuSerTyLeuThrIleValAlaProLeuValProProGlyIleThrProValIleAsn 480
Db 1381 ACCATATCTACGTTACTCAAAAGGACACTCTTGTAACAGCGGAGACCTGTGGTGAAC 1440
Qy 481 AlaLeuSerIleGluArgAlaMetLeuGluAsnIleMetArgAlaCyValGlyIleuAla 500
Db 1441 GCTTATCTTAAGACAGGCTATGCTGAGAAACATCTCAAGGCTGTGGTGGCTTGG 1500
Qy 501 ProGluAsnAsnMetIleLeuGluTyTyLys 510
Db 1501 CCAAGAACACATATGATCTTGGAAATCAAG 1530

RESULT 14
AF433879 1986 bp mRNA linear PIN 07-NOV-2001
LOCUS Sueda maritima subsp. sales myo-inositol-1-phosphate synthase
DEFINITION (INPS) mRNA, complete cds.
ACCESSION AF433879
VERSION AF433879.1 GI:16755888
KEYWORDS
SOURCE Sueda sales
ORGANISM Sueda sales
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Sueda.
REFERENCE
AUTHORS Wang,J.P., Wang,P.P., Sun,Y.F., Zhao,Y.X. and Zhang,H.
TITLE 1 (bases 1 to 1986)
JOURNAL Direct Submission
Submitted (15-OCT-2001) The Biology Department of Shandong Normal
University, Key Laboratory of Plant Stress Research, No.88, Wenhua
East Road, Jinan, Shandong Province 250014, China
FEATURES
source
1..1986
/organism="Sueda sales"
/mol_type="mRNA"
/sub_species="sales"
/db_xref="taxon:126914"
1..1986
/gene="INPS"
140..1672
/gene="INPS"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAI28131.1"
/db_xref="GI:16755889"
/translation="MPIENPKYSPNVNTENEHSVYVYETTELAVHNRKSGOWI
VAPKTVKQTHVFKLVGLVGVGVGNNCTLTGGVIANRBSGSMATKDIQANY
FESTLQASSIRVGSFNGEELVAPFKSLPMVNPDDI VFGWPISDLNADATRAKLV
DIDLQKLPVMEHWPLPGIYDPFIANODSRANNI IKTKKSCVEQVIDIREFK
EKKNYKVVVLTANTERSDYVGVNDTENTLMAVERSEI SPTLYAIAQED
VPFNGSPQNTFPGILIELAIKNCLIGDDRSQTKKSVLVDPIVAGAKPTISIV
SVNHGNDNGWLSAPQTRSEIKSVNVDMVANSGLVYRSGHPDVVVYKVPY
VSDSKRANDEYSEIFMGSKNTVLHNTEDSLAAPTILDLVLAESTRIEFALET
EDKFTFHPVATILSYLTKAPLVPRPTPVNALSKORAMLEKRAVCVGLAPENNMLE
YK"
ORIGIN
Alignment Scores:
Pred. No.: 7,12e-182 Length: 1986
Score: 2392.00 Matches: 460
Percent Similarity: 94.72% Conservative: 24
Best Local Similarity: 90.02% Mismatches: 25
Query Match: 90.88% Indels: 2
Gaps: 2
US-10-718-952-2 (1-510) x AF433879 (1-1986)
QY 1 MetPheIleGluAsnPhelYsValGluCysProAsnValIyetrThrGluThrgluIle 20
Db 140 ATGTTTATGCAAAATTTCAGGTGAGAAAGCCCAAACTTTACTACTGAAATGAAATT 199
QY 21 GlnSerValTyAsnTyrgIuThrThrgIuLeuValIleGluAsnArgAnslYThr--- 39
Db 200 CATTGAGTTTACATATGAAACCACTGATTAGTTTACAGAAATGCAAAAGGCTGTGCT 259
QY 40 TyrgIntPrlleValIySerProIySerValIyTyrgIuPhelYsThrAnIleHsVal 59
Db 260 TATCATATGATGCTCAAACTCAAAACGTCATATACCAATTCAAAACCTCAATGTT 319
QY 60 ProIySleuGlyValMetLeuValGlyTyrgIyAsnAnslYSerThrLeuThrgIy 79
Db 320 CCCAAACTTGGGGGTATGTTGGGATGGGGGTGGCAACACGTTGCCACTCACCGGT 379
QY 80 GlyValIleAlaAsnArgIuGlyIleSerTPrAlaThrIyAspIySileGlnAla 99

Db 380 GGTGTCATCGCCATCGGAGGGTATATCATGGGCAAAAGACAAATTCACACACT 439
QY 100 AsnTyrPheGlySerLeuThrGlnIleSerAlaIleArgValGlySerPheGlnIyGlu 119
Db 440 AACTATTTTGGATCCCTCACTCAGGCTTCTTATTCGTGTTGTTCTTTAATGGAGAG 499
QY 120 GlnIleTyraIaProPhelYsSerIleuLeuProMetValIaenProAspApIleValIhe 139
Db 500 GAGATTATGCTCCTTTAAGAGCCTTCTCTATGCGAACCAGATGATATAGTGT 559
QY 140 GlyIyTTPApIleSerAsnMetAsnIleuAlaAspAlaMetAlaArgAlIyValIhe 159
Db 560 GGGGGCTGGACATGAAGACTTGAACCTACAGACGCGCATGACGAGGCGAAAGGTGTTG 619
QY 160 AspIleAspLeuGlnIyGlnIleuArgProTyrmGluSerMetLeuProIeuProGly 179
Db 620 GACATTGATTTGCAAAAGACGTTGAGACCTTACATGAGACATGGTCCCTCTCTGT 679
QY 180 IleTyraAspProAspPheIleAlaIlaAnslngIuArgAlaAsnAsnValIleIyS 199
Db 680 ATCTATGACCCGTGATTGCTGCTCTAACCAAGATTCCTCGCAATATACATCATTAAG 739
QY 200 GlyThrIyGlnIuGlnIyValGlnIleIleIyAspIleIyValIphelYsGluAla 219
Db 740 GGCACCTAAGAGAGCAAGTGGAGCAAGTCATTAAGACATCAGGAAATTAAGAGAGAG 799
QY 220 ThrIySValAspIySValIValIleuThrThraIaenThrGluArgTySerAnIeu 239
Db 800 AACCAAGGTGACACAGGTGTGTACTCTGACAGCAAAACAGAGAGGTACATGATGTG 859
QY 240 ValValIyGlyLeuAsnAspThrMetGluAsnIleuValaIaValaAspArgAnsluAla 259
Db 860 GTGGTGGGACTCAATGACACCAACAGAACTTATGGCAGCTGTGAAGAAGATGACT 919
QY 260 GlnIleSerProSerThrIeuTyraIleAlaIyCysValMetGluAsnValProhelle 279
Db 920 GAGATTTTCTCAATCACTTGTATGCTATGCTTGCACCAAGAAAGTCCCTTCAATC 979
QY 280 AsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAn 299
Db 980 AACGGTATGCTTCAGAACACCTTGTCTCTGTCTCATTTGAACGTGCCATTAAGAAAGAT 1039
QY 300 ThrIleuIleGlyIyAspAspPheIySergIyGlnThrIyMetIySerValIeuVal 319
Db 1040 TGTTTGATTGGAGGAGATGACTTCAAGAGTGTGACATAAGATTAATCACTGCTCGTT 1099
QY 320 AspPheIeuValGlyIyAlaGlyIleIySProThrSerIleValSerTyraAsnHsIeuGly 339
Db 1100 GACTTCCCTTGTGGGGCTGGTATTAAGCCACATCAATTTGTAGCTAACCACTTGGGA 1159
QY 340 AsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIyGlnIleSerIyS 359
Db 1160 AACCAATGATGGATAACTTTCAGACCTTCAGACCTTCAGGTCTTAAGAGATTTCAAAAG 1219
QY 360 SerAsnValIaIaAspAspMetValaAsnSerAsnAlaIleIeuTyrgIuProGlyGluHs 379
Db 1220 AGTATATGTGTTGATGACATGTTGCTAGCAACGCTATCTGTATGACCTGTGTGAACAC 1279
QY 380 ProAspHsIvalIaIyValIleIySfyTyraIProTyTyraIyAspSerIyArgAlaMet 399
Db 1280 CCTGATCATGTTGTTGTATCATCAAGTATGTCATATGTCGAGAGACGAAAGAGACTATG 1339
QY 400 AspGluTyrrThrSerGlnIlePheMetGlyIyIySserThrIleValIleuHsIaenThr 419
Db 1340 GATGAGTACATCTGAGATATTCATGGGTGATGAACACCATGTTTGGCAACACACC 1399
QY 420 CysGluAspSerIleuAlaIaProIleIleuAspLeuValIleuAlaIuIeu 439
Db 1400 TGTGAGGACTGCTCTTGGCGGCTCCATCATCTTGAAGCTTGCTCTTGGCGAGCTT 1459
QY 440 SerThraGlyIleIuPhelYsAlaGluAsnIyGlyIyPheHsSerPheHsIaProVal 459
Db 1460 AGCACCCGATTAATTCAAAGCCGAAACCTGAGGACAAAGTTCCACACCTTCATCCAGT 1519

RESULT 15
MCU32511
LOCUS MCU32511
DEFINITION Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase mRNA, complete cds.
ACCESSION U32511
VERSION U32511.1 GI:975887
KEYWORDS
SOURCE
ORGANISM Mesembryanthemum crystallinum (common iceplant)
REFERENCE Mesembryanthemum crystallinum
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Aizoaceae; Mesembryanthemum.
TITLE 1 (bases 1 to 2053)
JOURNAL Ishitani,M., Majumder,A.L., Bornhauer,A., Michalowski,C.B.,
MEDLINE Jensen,R.G. and Bohner,H.J.
REFERENCE Coordinate transcriptional induction of myo-inositol metabolism during environmental stress
AUTHORS Plant J. 9 (4), 537-548 (1996)
TITLE 2 (bases 1 to 2053)
JOURNAL Ishitani,M., Majumder,A.L., Bornhauer,A., Michalowski,C.B.,
MEDLINE Jensen,R.G. and Bohner,H.J.
REFERENCE Direct Submision
TITLE Submitted (27-JUL-1995) Hans J. Bohner, Biochemistry, University
JOURNAL of Arizona, Biochemistry West 516, Tucson, AZ 85721, USA
FEATURES
source Location/Qualifiers
1..2053
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
182..1720
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAB03687.1"
/db_xref="GI:975888"
/translation="MFIESPVSBNVKTENIESVYNYDTTELVHNRKDGQYOM
IVKTYOVHFKTDTRVPKCYVMFKSLGNGNGSTLGGVIANRGISWARKDKIKQAN
YRSLTQASSIRVGSFNGEBIYAPFGKSLPMVNDVDFGMDISNDNLADAMTRAV
FDIDLOQLRPMEHMYPLPGIYDPDTIANQSSRANVVKGTKEQVERVADIREF
KEKKRVKVVVLTGNTERSNVVGVANDTENTLALERKESISISTVALACIEE
NIPIINSPOTFVPGILDLAKNGSLIGDDFKSGQTKMSVVDLVGAGIKPTSI
VSYVHLGNNDGNLSAPOTFPSKESISNVVDMVAVDILYEGEHPDHYVYKYVP
VYSGSKRAMDEYSEIFMGCGNTIVGNTDSDSLAIIIDLVILAEISRTOLKAE
BEDRHSFHPVATILSYTLAPLVPGTPVYNALSKORAMENILRACVGLAPENNNT
LETK"
ORIGIN
Alignment Scores:
Pred. No.: 5,63e-181 Length: 2053
Score: 2381.00 Matches: 454
Percent Similarity: 94.73% Conservative: 31
Best Local Similarity: 88.67% Mismatch: 25
Query Match: 90.46% Indels: 2
Gaps: 1
US-10-718-952-2 (1-510) x MCU32511 (1-2053)

Db	182	ATGTTTATTGAGAGCTTCAAGGTACAGAGCCCAACGTAAGTACACAGAAATGAGATT	241
Oy	21	GlnSerValTyrAsnTyrGluThrThrgluLeuValHisgluAsnArgAsn-----Gly	38
Db	242	GAATCGGATGACACTACAGACACACGAGTGGTTTCATGGAATCGCAAGATGCTGCT	301
Oy	39	ThrTyrGlnTrpIleValLysProLysSerValLysTyrGluPheLysTrpAsnIleHis	58
Db	302	GGTTATCATAGTGAATCGTTAAAGCCCAAGACCTGTACATCATCTTCAAGACTGACATCTGT	361
Oy	59	ValProLysLeuGlyValMetLeuValGlyTrpGlyGlyAsnAsnGlySerThrLeuThr	78
Db	362	GTCCCAAACTCGGAGGATGCTGTTGGATGGGAGTGAATAATACGGCTCCACCTTCA	421
Oy	79	GlyGlyValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGln	98
Db	422	GGTGGTGTGATCGGCACACCGGAGGAAATCTTGGCCACCAAGACAAATATTCACAGAA	481
Oy	99	AlaAsnTyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGly	118
Db	482	GCGAACATATTTGGTTCGTGATGCCACGAGCTTCTTCAATTCGGGCTTTTAAATGGA	541
Oy	119	GluGluIleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleVal	138
Db	542	GAGAGATTTAATGTCTTCTTCMAAGCTGTCTCCCTTGTGGAAACCAAGATGACTGTG	601
Oy	139	PheGlyGlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysVal	158
Db	602	TTTGGGGGTTGGGACATPAGGACATGAACTGCTGTATGCAATGACCAAGGCTTACGGCTC	661
Oy	159	PheAspIleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuPro	178
Db	662	TTTGACATATGATCTTCAAAAGCAGCTCAGGCTTACATGGAGCACAATGGTTCCCTCCT	721
Oy	179	GlyIleTyrAspProAspPheIleAlaAlaAsnGlnGluGluGlyAlaAsnAsnValIle	198
Db	722	GGATTTTACAGACCTGTATTCATGTGCTGCGCAACAAAGCTCCCAAGCCAAACAGTCA	781
Oy	199	LysGlyThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlu	218
Db	782	AMGGGACCAAGAAAGAAACAGATTGAGCGAGCTATTAAAGCATTTAGGAGCTTCAAGAG	841
Oy	219	AlaThrLysValAspLysValAlaValIleuTrpThrAlaAsnTrpGluArgTyrSerAsn	238
Db	842	AAAGAACAAAGTGACAAAGTGCTGTCTATGACTGGCAACACAGAGAGGTACAGCAAT	901
Oy	239	LeuValValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValaAspArgAsnGlu	258
Db	902	GTGTGTGTGTGGGCTCAACGACACCATGAGAACTGTGGCATCTTCGAGAAATAAGAG	961
Oy	259	AlaGluIleSerProSerThrLysTyrAlaIleLeuLysValMetGluAsnValProPhe	278
Db	962	TGGAGATTTCTTCATCATCTGATGCTGTATGCTTGCATGATGCAAGAAATAATTCCTTC	1021
Oy	279	IleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArg	298
Db	1022	ATCATATGATGAGCCTTCAGAACACTTGTGTGCGAGGCTCATTTGATCTGGCTATCAAGAG	1081
Oy	299	AsnThrLeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeu	318
Db	1082	AACAGTTTGAATGTGTGTGATGACTTCAAAAGTGTGACAGCAACAAGAAATGCTGCTTC	1141
Oy	319	ValAspPheLeuValGlyAlaGlyIleLysProThrSerIleValaSerTyrAsnHisLeu	338
Db	1142	GTCACTTCCCTCGTGGGCTGGATTAACCCAAACATCCATAGTACCTTCAACACACTTG	1201
Oy	339	GlyAsnAsnAspGlyMetCAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSer	358
Db	1202	GGGACAAATATGAGAGTGAACCTGTCTGACACTCTTCAACTTTCAGTCTCAAGAGATTTCT	1261
Oy	359	LysSerAsnValValaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlu	378

Db 1262 AAGAGCAACGTGATGATGATGTTGCCAGCAACGGCATCCTCATGAGCCGGGCGAA 1321
QY 379 HisProAspHisValValValIleIysTyrValProTyrValGIYAspSerLysArgAla 398
Db 1322 CACCTGATCATGTTGTGATCATCAAGTATGTGCCATATGTGGAGACACGACAGAGAGCC 1381
QY 399 MetAspGluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsn 418
Db 1382 ATGATGAGTACACTTCGGAATAATTATGGGGGAGAACGACACTATGTGCATGCAACAC 1441
QY 419 ThrCysGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValleuLeuAlaGlu 438
Db 1442 ACCGTGAGAGACTCTCTGTGGCTGCCCCCATCATCTGGACTTGGTCTCTGCTGNA 1501
QY 439 LeuSerThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisPro 458
Db 1502 CTAAAGACCCGCACTTCAGCTGAAGGCTGAGAGAGAGACAAAGTTCCACTCTTCCATCCG 1561
QY 459 ValAlaThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProVal 478
Db 1562 GTCGCTACCATCTCTGAGTTAAGCTCACCAAGGCCCTCTTGTCCACAGGCACTCCAGTG 1621
QY 479 ValAsnAlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGly 498
Db 1622 GTPAATGCACTTCGAAGCAGAGGCTATGCTGAGAACATACATACTGAGGGCTGTGTGGT 1681
QY 499 LeuAlaProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1682 TTGGCTCTGAGAACACATGATCTCGAATACAG 1717

Search completed: June 7, 2005, 20:45:00
Job time : 5362 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:10:41 ; Search time 633.5 Seconds
(without alignments)
4765.695 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MLENPKVCECPVKTETETI.....NIMRACVGLAPENMILEYK 510

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=x1p
-O=/cgn2.1/USPTO.spool.p/US10718952/runat.06062005.173400.12840/app_query.fasta_1.1294
-DB=N Geneseq.16Dec04 -QFMT=fasta -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-NODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952.0CGN.1.1.886.0runat.06062005.173400.12840 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: N_Geneseq.16Dec04:*
2: geneseqn1908:*
3: geneseqn1908:*
4: geneseqn2000a:*
5: geneseqn2001a:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2632	100.0	1533	12	ADQ14498
2	2632	100.0	1533	13	ADQ14498
3	2632	100.0	1533	13	ADQ14498
4	2632	100.0	1760	12	ADQ14490
5	2632	100.0	1782	2	AAV62440

6	2627	99.8	1533	2	AAV62443	AAV62443 Soybean m
7	2627	99.8	1533	12	ADQ14494	AdQ14494 Mutant so
8	2627	99.8	1533	13	ADQ14494	AdQ14494 Mutant so
9	2600	98.8	1533	12	ADQ14504	AdQ14504 Wild type
10	2600	98.8	1533	13	ADQ14502	AdQ14502 Mutant so
11	2600	98.8	1533	12	ADQ14502	AdQ14502 Mutant so
12	2600	98.8	1533	13	ADQ14502	AdQ14502 Mutant so
13	2593	98.5	1533	12	ADQ14500	AdQ14500 Mutant so
14	2593	98.5	1533	13	ADQ14500	AdQ14500 Mutant so
15	2472	93.9	1530	2	AAV62402	AAV62402 Mutant so
16	2378	90.3	1533	6	AB213633	Ab213633 Arabidops
17	2378	90.3	1533	12	ADN73524	Adn73524 Thale cre
18	2378	90.3	1533	3	AAV62422	AAV62422 Arabidops
19	2368	90.0	1781	4	AAV62422	AAV62422 Arabidops
20	2352.5	89.4	1536	3	AAV62428	AAV62428 Arabidops
21	2352.5	89.4	1536	3	AAV62428	AAV62428 Arabidops
22	2346	89.1	1665	2	AAV62406	AAV62406 Arabidops
23	2336	88.8	1931	2	AAV62407	AAV62407 Inducible
24	2335	88.7	1959	4	AAV62407	AAV62407 MIP synth
25	2283	86.7	1759	6	ABQ72653	Abq72653 Human MDD
26	2129.5	80.9	1719	3	AAV62405	AAV62405 Arabidops
27	2026	77.0	1536	12	ADP43918	Adp43918 P. coatc
28	1813	68.9	3546	2	AAV62441	AAV62441 Maize myo
29	1812	68.8	3546	2	AAV62441	AAV62441 Maize myo
30	1607	61.1	1772	4	AAH14171	Aah14171 Human CDV
31	1607	61.1	1818	4	AAH15251	Aah15251 Human CDV
32	1607	61.1	1825	6	AAH17176	Aah17176 Human ova
33	1607	61.1	1852	4	AAH17741	Aah17741 Human hml
34	1607	61.1	1862	4	AAH19232	Aah19232 Human pol
35	1607	61.1	2380	13	ADQ86789	Adq86789 Human tum
36	1607	61.1	2380	13	ADQ86789	Adq86789 Human tum
37	1607	61.1	2380	13	ACN40520	Acn40520 Tumour-as
38	1596.5	60.7	1605	8	ABT19345	Abt19345 Aspergill
39	1593	60.5	1833	4	AAH47740	Aah47740 Murine m
40	1592.5	60.5	1833	4	AAH47740	Aah47740 Murine m
41	1585	60.2	2280	8	ABT21165	Abt21165 Aspergill
42	1556.5	59.1	1980	12	ADQ00499	Adq00499 Novel hum
43	1556.5	59.1	1980	12	ADQ00499	Adq00499 Novel hum
44	1547	58.8	1991	12	ADG73735	Adg73735 Aspergill
45	1530	58.1	1704	8	ABT20567	Abt20567 Aspergill

ALIGNMENTS

RESULT 1	ADQ14498	ADQ14498 standard; cDNA; 1533 BP.
ID	ADQ14498	ADQ14498
XX	ADQ14498	ADQ14498
AC	ADQ14498	ADQ14498
XX	ADQ14498	ADQ14498
DT	23-SEP-2004	(first entry)
XX	23-SEP-2004	(first entry)
DE	Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.	
KW	Soybean; myo-inositol 1-phosphate synthase; gene; ss;	
KW	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;	
KW	inositol phosphate; mutant.	
XX	Glycine max.	
OS	Synthetic.	
XX	Synthetic.	
FT	Key	Location/Qualifiers
FT	CDS	1..1533
FT	FT	/product= "Mutant soybean myo-inositol 1-phosphate synthase #2"
PN	US2004128713-A1.	
XX	US2004128713-A1.	
XX	01-JUL-2004.	
PD	01-JUL-2004.	
XX	01-JUL-2004.	
PF	21-NOV-2003; 2003US-00718952.	

PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-0506822.
PR 26-APR-1998; 98US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX (HIRTZ/) HIRTZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hirtz WD, Sebastian SA, Grace DJ, Streit LG;
DR WPI; 2004-533135/51.
XX P-PSDB; ADQ14499.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Example 8; SEQ ID NO 9; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,98e-253	Length:	1533
Score:	2632.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-718-952-2 (1-510) x ADQ14498 (1-1533)

QY 1 MetPheIleGluAsnPhelYsValGluCysProAsnValIysTyRThrgIuThrgIuile 20
DB 1 AAGTTCATCGAAGATTTTAAGGTGTGAGTGTCTTAATGTGAAGTACCCAGACTGAGATT 60
QY 21 GlnSerValIysAsnTyRgluThrThrgIuLeuValIleGluAsnArgAsnGlyThrTyR 40
DB 61 CAGTCTCGTGTACAACTACGAAACACCGAAGCTGTTCACGAGAACGGAATGGCACTTAT 120
QY 41 GlnTrpIleValIysProIysSerValIysTyRGlulPhelYsThraSnIleHisValPro 60
DB 121 CAGTGAATGTTCAAACCCAAATCTGTCAATACGAAATTTAAACCAACATCTCATGTCTT 180
QY 61 LysLeuGluValMetLeuValGlyTyRGlYAsnAsnGlySerThrLeuThrgIyGly 80
DB 181 AATATTGGGGTAAATGCTTGTTGGGTGGGAGAAACAACGGCTCAACCCCTCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnIleAlaAsn 100
DB 241 GTTATTGTCAACGAGAGGGCATTTATGCGCTACAAAGACAAAGATTCACCAAGCAAT 300

QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
DB 301 TACTTTGGCTCCCTCACCAGAGCTCAGACTATCCGAGTTGGGTCTTCCAGGAGAGAA 360
QY 121 IleTyRAlaProPhelYsSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTCMAAGAGCTGCTTCMAATGGTTAAACCTCGAAGACATGTGTGGG 420
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
DB 421 GGATGGATATCAGCAACATGAACTGGCTGATGCCATGGCCAGGCAAGAGTGTTCAC 480
QY 161 IleAspLeuGlnIysGlnLeuArgProTyRMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTGCAGAAACAGTGTAGAGCTTACATGAAATCATGCTTCCACTCCCGGAATC 540
QY 181 TyRAspProAspPheIleAlaIleAsnGlnGluValGlnAlaAsnValIleIleGly 200
DB 541 TATGACCCGGATTTCTATTGCTGCCAACCAAGAGCGCTGCCAACAGTCATCAAGAGC 600
QY 201 ThrIysGlnGluGlnValGlnGlnIleIleYsAspIleYsAlaPheIysGlnAlaThr 220
DB 601 ACAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGCGCTTTAAGAGCACCC 660
QY 221 LysValAspLysValValIleuThrPhrAlaAsnThrgIuArgTyRSerAsnLeuVal 240
DB 661 AAAGTGCACAAAGTGTGTACTGTGACTGCTGCACACAGAGAGTACAGTATTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaIleValAspArgAsnGluAlaGlu 260
DB 721 GGGGCGCTTAAAGACACATGAGATCTCTTGCTGTGTGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyRAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCCACCTGTATGCCATTCGTTGTGTGAAGAAATGTTCTTCAATTAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATTGATCTTGCCATCGCGAGAACT 900
QY 301 LeuIleGlyIleAspAspPheIysSerGlyGlnTrpIysMetIysSerValLeuValAsp 320
DB 901 TTGATTGGTGGAGAGACTTCMAAGTGTGTGACCAAAATGAAATCTGTGTGTAT 960
QY 321 PheLeuValGlyAlaGlyIleYsProThrSerIleValSerTyRAsnIleGluIysAsn 340
DB 961 TTCCTTGGGGGGCTGGTATCAGCCCAACATCATATAGCTTACACCATCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIysGluIleSerIysSer 360
DB 1021 AATGATGGTAAATCTTGGCTCCCAAACTTCCGTTCCAAAGAAATCTCCAAAGAGC 1080
QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyRGlulProGlyGluIlePro 380
DB 1081 AACGTTGTGATGATATGATGTCAACAGCAATGCCATCTTAAGAGCCGTGGTAAATCCA 1140
QY 381 AspHisValIleValIleYsTyRValProTyRValGlyAspSerIysArgAlaMetAsp 400
DB 1141 GACCAATGTTGTATTAAGATATGATGCTTACGTAAGGAGACAGAAAGACCATGAT 1200
QY 401 GlyTyRThrSerGluIlePheMetGlyIysSerThrIleValIleHisAsnThrCys 420
DB 1201 GAGTACACTCAGAAATATTCATGGGTGGAAAGACACACATGTTTCCACAAACATGC 1260
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTCTGCTGCTCTTAATCTTGAGCTTGCTCTTCTTGCTAGCTCAGC 1320
QY 441 ThrArgIleGluPhaIysAlaGluAsnGluGlyLysPheHisSerPheHisIleProValAla 460
DB 1321 ACTAAGATCGAATTTAAAGTGAAGAAATGAGGAAATTCATCTCATCCACCAAGTGTCT 1380
QY 461 ThrIleLeuSerTyRLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480

DB 1381 ACCATCTCAGCTACTCACCAGGCTCCTCTGTTCCACCGGGTACACCAAGTGGAT 1440
QY 481 A1A1eSeLryGlnAryA1AmeLleuGluAenIlIeWeArA1AcYva1GlyLeuA1a 500
DB 1441 GCATTCTCAAGCAGCGCTCAATGCTGGAAACATATAGAGGCTGTGTGATGGCC 1500
QY 501 ProGluAenAmeLleuGluYrlye 510
DB 1501 CCAGAGATTAACATGATTCTCGATCAAG 1530
RESULT 2
ADS81999
ID ADS81999 standard; cDNA; 1533 BP.
XX
AC ADS81999;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29004JP01.
XX
FH Key Location/Qualifiers
XX CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
PI Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
DR P-PSDB; ADS82000.
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX
PS Example 8; SEQ ID NO 9; 34p; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the syntheses for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a

CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.98e-253 Length: 1533
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Gaps: 0
DB: 13 Indels: 0
US-10-718-952-2 (1-510) x ADS81999 (1-1533)
QY 1 MetPheLleGluAenPheLyseValGluCyseProAenValIySeThrGluThrGluIle 20
DB 1 ATGTCATCGAGAAATTTTAAGGTAGTGTCTTAATGTGAATACCCAGACTGAGATT 60
QY 21 GlnSerValIyYAsnIyTyrGluThrGluLeuValHlGluAenAryAenIyThrIyYr 40
DB 61 CAGTCGCTGTACAACTACGAAACACACGAACTGTTCACGAAACAGAAATGGACCTAT 120
QY 41 GlnTrpIleValIyProIySerValIySeryIyGluPheLyseThrAenIleHlValPro 60
DB 121 CAGTCGATGTCAAACCCCAATCTGTCAAAATCGAATTTAAACCAACATCCATGTTCT 180
QY 61 LysLeuGluValIleLeuValGlyTrpGlyAenAenGlySerThreuthrGlyGly 80
DB 181 AAATTAGGGGTATATCTGTGGGTGGGTAACAACAGGCTCAACCTCACCGGTGGT 240
QY 81 ValIleA1AenAryGluGlyIleSerTrpA1AenIyAenIyAenIyGlnIleA1Aen 100
DB 241 GTTATTGCTAACCGAGAGGCACTTCATGCGCTACAAAGACAAATTCACAGCGAAT 300
QY 101 TyrPheGlySerLeuThrGlnA1AseRajleAryValGlySerPheGlnIyGluGlu 120
DB 301 TACTTGGCTCCCTCAACCAAGCCTCAGCTATCCGAGTTGGGTCTTCAGAGGAGGAA 360
QY 121 IleTyrA1AenPheLyseThreuthrPheLeuProMetValAenProAenAenIleValPheGly 140
DB 361 ACTATAGCCCCCAATTAAGAGCCTGTCTTCATATGTTAACCTCGAGACATTTGTGTTGG 420
QY 141 GlyTPAeP1LeSeRAsmeLleuA1AeP1AmeRajleAryA1AlyseValPheAeP 160
DB 421 GGATGGATATACGAACATGAACCTGGCTATGCTACAGGCAAGAGTGTGTTGAC 480
QY 161 IleAePLeuGlnIyGlnLeuAryProIyMetGlnIySerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTGGCAAGACAGATTGAGGCTTACATGAATCATCTCTCCATCCCGGAATTC 540
QY 181 TyrAePProAePheIleA1A1AenGlnIyGluA1AenAenValIleIyGly 200
DB 541 TATGACCCGGAATTTATGCTGCAACAAAGAGGCGTCCAAACAGCTCAAGAGGC 600
QY 201 ThrIySGlnGluGlnValGlnGlnIleIleIyAeP1IleYsA1A1PheIySGlnA1A1Thr 220
DB 601 ACAAAGCAAGACCAAGCTTCAACAAATCATCAAGACATCAAGAGCTTTAAGAAACCAAC 660
QY 221 IySValAePLeuValIleValIleuTrpThra1AenThrGluArySerAenLeuVal 240
DB 661 AAAGTGACAAAGGTGTGTACTGTGACTGCAACACAGAGAGGTACATGTAATTTGGTT 720
QY 241 ValGlyLeuAenAePThrMetGluAenIleuA1A1A1A1A1A1A1A1A1A1A1A1A1A1 260

```

Db      |||||
721 GTGGGCTTAATGACCATGAGAAATCTTGGCTGTGAGACAGAAATAGAGCTGAG 780
Qy      |||||
261 IIESerProSerThreutyrAlaIleAlaCysValMetGluAsnValProheleAsn 280
Db      |||||
781 ATTTCTCTTCCACCTTGATGACCATGCTGTGTATGAAAAGTTCTCTTCAATTAAT 840
Qy      |||||
281 G1ySerProG1AsnThrPheValProG1yLeu1eAspLeuAla1leAlaArgAsnThr 300
Db      |||||
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTGATCTTGCCATCGCGAGAACCT 900
Qy      |||||
301 Leu1leG1yG1yAspAspPhe1ySerG1yG1nThr1yMet1ySerVal1leuValAsp 320
Db      |||||
901 TTGATTGGGGAGAGACACTTCAAGAGTGGTCAGACCAAAATCTGTGTGGTGTAT 960
Qy      |||||
321 PheLeuValG1yAlaG1y1le1ySerProThrSer1leValSer1yAsn1leuG1yAsn 340
Db      |||||
961 TTCTCTGTGGGGGCTGTGTATCAAGCCACATCTATAGTACATTACACCATCTGGGAAAC 1020
Qy      |||||
341 AenAspG1yMetAsnLeuSerAlaProG1nThrPheArgSer1ySerG1u1leSer1ySer 360
Db      |||||
1021 AATGATGATGATGATTTTCCGCTCCACAACTTCCGTTCCAGGAATCTCCAGAAC 1080
Qy      |||||
361 AenValValAspAspMetValAsnSerAsnAla1leLeu1yG1uProG1yG1uHisPro 380
Db      |||||
1081 AACGTTGTTGATGATGATGTCACACAGCAATGCCATCTCTATGAGGCTGTGACATCCA 1140
Qy      |||||
381 AspHisValVal1le1ySer1yValPro1yValG1yAspSer1yAsnArg1leMetAsp 400
Db      |||||
1141 GACCATGTTGTGTATTAAGATGTCCTTACGTAGGGGACAGCAAGAGCCATGAT 1200
Qy      |||||
401 G1u1yThrSerG1u1lePheMetG1yG1y1ySerThr1leVal1leuHisAsnThrCys 420
Db      |||||
1201 GAGTACACTTCAGAGATATTCATGGGTGAAAAGACACATCTTTTGACAAACATCTC 1260
Qy      |||||
421 GluAspSer1leu1leAlaPro1le1leLeuAsp1leu1leu1leu1leu1leuSer 440
Db      |||||
1261 GAGGATTCCTCTTACGCTCTCTCTATTAATCTTGACATCTGCTTCTCTGACGCTCAGC 1320
Qy      |||||
441 ThrArg1leG1uPhe1yValG1uAsnG1uG1y1ySerPheHisSerPheHisProValAla 460
Db      |||||
1321 ACTAGATGAGATTAAAGCTGAAAATGAGGAAAATTCACATCTTCCACCCAGTTGCT 1380
Qy      |||||
461 Thr1leLeuSer1yLeuThr1yAsn1yAsnProLeuValProProG1yThrProVal1leAsn 480
Db      |||||
1381 ACCATCTCAGCTACTCTACCAAGGCTCCTCTGTTCCACCGGTATCACACATGCGTGAAT 1440
Qy      |||||
481 AlaLeuSer1yG1nArg1leMetLeuG1uAsn1leMetArg1leCysValG1yLeuAla 500
Db      |||||
1441 GCATTGTCAAAGCAGCGTCAATGCTCGGAAAACATTAATGAGGCTTGTGTGATTGGCC 1500
Qy      |||||
501 ProG1uAsnAsnMet1leLeuG1u1y1y1y 510
Db      |||||
1501 CCAGGAATTAACATGATCTTCGAGTACAAAG 1530

```

RESULT 3
AD881993
ID AD881993 standard; cDNA; 1533 BP.

AC AD881993;
XX
XX 18-NOV-2004 (first entry)
DT
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; cultivar Wye.
OS
XX
XX Key Location/Qualifiers
FH 1.1533
FT CDS /tag= a
FT

```

FT      /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
PN
XX
XX 17-APR-2003.
PD
XX
XX 11-MAR-2002; 2002US-00025003.
PF
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
PI
XX
XX MPI: 2004-639957/62.
DR P-PsDB; AD881994.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 2; SEQ ID NO 1; 34pp; English.
PS
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences), where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.
XX
XX
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	1,986-253	Length:	1533
Score:	2632.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-10-718-952-2 (1-510) x AD881993 (1-1533)

```

Qy      |||||
1 MetPhe1leG1uAsnPhe1ySerValG1uCysProAsnVal1y1yThrG1uThrG1u1le 20
Db      |||||
1 ATGTTCAATCGAAGATTTTAAAGTTAGCTGTCTATGTGAATACCCAGAGCTAGATTT 60
Qy      |||||
21 G1nSerVal1yAsn1yG1uThr1yThrG1uLeuVal1leG1uAsnArgAsnG1yThr1y1y 40

```

```

Db 61 CAGTCGGTGTACAACTACGAAACCGAAGCTGTTCACGAGAAACAGGATGGCACCTAT 120
Qy 41 GATTTTLeValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db 121 CAGTGAATGTCAAAACCAAAATCTGTCAAAATAGCAATTTTAAACCAACATCCATGTCTCT 180
Qy 61 LysLeuGlyValMetLeuValGlyTyrPGLyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AAATTGGGGTAAATGCTGTGGGTTGGGGTGGAAACAAGGCTCAACCTCACCGGTGGT 240
Qy 81 ValIleAlaAsnArgGlyGlyIleSerTyrPalaThrLysAspLysIleGlnGlnAlaAsn 100
Db 241 GTATTATGCTAACCGAGAGGGCATTTTCATGGCTTACAAAGAGCAAGATTCACAAAGCAAT 360
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTTTGGCTCCCTCCACCCCAAGGCTCAGCTATCCGAGTTGGTCTTCCAGGAGAGGAA 360
Qy 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTATAGCCCATTCACAGAGCTGTCTTCATAGGTTAACCTTACACGATGTGTGGG 420
Qy 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GGAATGGATATACGAACATGAACTGTGCTGATGCCATGGCCAGGCAAGGCTTTGAC 480
Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db 481 ATCGAATTTGACAGAGAGGTTGAGGCTTACATGAAATCCATGCTCCACCTCCCGGATC 540
Qy 181 TyrAspProAspPheIleIleAlaAsnGlnGluLysArgAlaAsnValIleLysGly 200
Db 541 TATGACCCGATTTCAATGTGTCACCAAGAGAGAGGTCACCAACGTCATCAAGAGGC 600
Qy 201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysValPheLysGlyAlaThr 220
Db 601 ACAAAACAGAGAGAGAGGTTCAACAAATCATMAAGACATCAAGGCTTTAAGAAAGCCACC 660
Qy 221 LysValAspLysValValIleLeuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 661 AAAGTGGACAAAGGTGTTGTATGTAAGTGCACCAACAGAGAGATACATTAATTTGGTT 720
Qy 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAlaAspArgAsnGluAlaGln 260
Db 721 GTGGGCTTATATACACCATGAGAAATCTTGTGCTGTGAGCAAGAAATGAGGCTGAG 780
Qy 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATTTCTCTTCCACCTGTATGCAATGCTGTGTGTATGGAATGTTCTTTCATTTAT 840
Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAGCCCTCAGAAACCTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGAAACACT 960
Qy 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTGGTGAATACCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTGTGAT 960
Qy 321 PheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 961 TTCCTTGTGGGGCTGTATCAAGCCAACTCTATATGTCATTTACAAACCACTTGGGAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db 1021 AATGATGTATGTAATCTTGTGGCTCCACAACTTTCGTTCCAAAGAAATCTCCAAAGAG 1080
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db 1081 AAGGTTGTGATATGTGTCAACAGCAACATGCTCTATATAGGCTGTGTGAACATCCA 1140
Qy 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400

```

```

Db 1141 GACCATGTTGTTGTATATATAGTATGCTTACAGAGGAGACAGAGAGCCATGAT 1200
Qy 401 GlnTyrThrSerGlnIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db 1201 GAGTACACTTCACAGATATTTCAATGGGTGGAAGAGCAACATTTGTCACACACATGC 1260
Qy 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGATTCCTCTTACTGCTGCTCTATATCTTGGACTTGGTCTTCTTGTGAGCTCAGC 1320
Qy 441 ThrArgIleGluPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAGATACGAGTTTAACTGAAATAGAGGAAATTCACATTCACCCAGATTGCT 1380
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATCTCAGCTACCTCCACCAAGGCTCTGTGTGTCCACCGGTACACAGAGTGAAT 1440
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCATTTGCAAGAGGCTGCAATGCTGGAACATATATAGGGCTGTGTGATTTGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db 1501 CCAGAGATATACATGATTTCTCGAGTACAAAG 1530

RESULT 4
ADQ14490
ID ADQ14490 standard; cDNA; 1760 BP.
XX
AC ADQ14490;
XX
DE 23-SEP-2004 (first entry)
XX
XX wild type soybean myo-inositol 1-phosphate synthase cDNA #1.
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
XX myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate.
XX
XX Glycine max.
XX
XX Key Location/Qualifiers
FH 54..1586
FT /tag= a
FT /product= "wild type soybean myo-inositol 1-phosphate
FT synthase #1"
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751..
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX MPI; 2004-533135/51.
XX DR P-PSDB; ADQ14491.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.

```

XX Claim 4: SEQ ID NO 1; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant,
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents cDNA encoding a wild type
 CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

SO Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,4e-253	Length:	1760
Score:	2632.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-718-952-2 (1-510) x ADQ14490 (1-1760)

QY	1	MetPheIleGluAsnPhelysValGluCyPProAsnValIleTyThrGluThrGluIle	20
DB	54	ATGTTTCATGAGAAATTTTAAGTGTGAGTGTCTTAAGTGAAGTACCCAGACTGAGATT	113
QY	21	GlnSerValIleAsnTyrgluThrThrGluLeuValHleGluAsnArgAsnGlyThrTy	40
DB	114	CAGTCGGTTCACACTACGAAACCCGAACTTTTTCACGAGAACGGAATGGCACTTAT	173
QY	41	GlnTrpIleValIleAsnProIleSerValIleTyrgluPheIleThrAsnIleHisVal	60
DB	174	CAGTGGATTGTCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCT	233
QY	61	IleLeuGluIleValMetLeuValGlyTrpGlyAsnAsnGlySerThrLeuThrGly	80
DB	234	AAATTAGGGGTAAATGCTTGTGGTGGGTGGAAACACGGCTCAACCTCTCACCGGTGT	293
QY	81	ValIleAlaAsnArgGluGlyIleSerTrpAlaThrIleAsnProIleGlnIleAlaAsn	100
DB	294	GTTATTGCTACCGAGGAGGCAATTTCAATGGCTACAAAGACAAAGTTCAACAGCAAT	353
QY	101	TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGlu	120
DB	354	TACTTGGCTCCCTCCCAAGCCTCAGCTATCCGAGTTGGGCTTCCAGGAGAGAA	413
QY	121	IleTyrAlaProPheIleSerLeuLeuProMetValAsnProAsnAspIleValPheGly	140
DB	414	ACTTAAGCCCATTCAGAGGCTGCTTCCAAATGTGTAACTTCAGCAACATGTGTGGG	473
QY	141	GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp	160
DB	474	GGATGGGATATTCGACAACTGAACCTGGCTGATGCCATGGCCAGGCAAAAGTGTGGAC	533
QY	161	IleAspLeuGlnIleGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle	180
DB	534	ATCGATTTCAGAGAGAGTTGAGGCTTTCATGGAATTCATGTTCCACTCCCGGAATC	593
QY	181	TyrAspProAsnPheIleAlaAsnGlnGluIleArgAlaAsnAsnValIleIleGly	200
DB	594	TATGACCCCGGATTTCTGCTGCAACCAAGAGACGTCGCAACACATCATCAAGGGG	653

QY	201	ThrIleGlnGluIleValGlnGlnIleIleIleAspIleIleValAlaPheIleGluIleAlaThr	220
DB	654	ACAAAGCAAGAGCAAGTTCACCAATCATCAAAAGCATCAAGAGCTTTAAGGAACCAAC	713
QY	221	LysValAspIleValIleValIleLeuTrpThrAlaAsnThrGluIleArgTyrSerIleLeuAl	240
DB	714	AAAGGAGCAAGGTGGTGTATGTGGACGTGCACACAGAGAGGTACAGTAATTTGGTT	773
QY	241	ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluIleGlu	260
DB	774	GTGGGCTTATGACACATGAGAGATCTTTGGCTGCTGTGCACAGAAATAGGCTGAG	833
QY	261	IleSerProSerThrLeuTyrAlaIleAlaCyValMetGluAsnValProPheIleAsn	280
DB	834	ATTCTCTCCATCACTTGAATCCATGCTGTGTGTATGAAATAGTTCTTTCATTAAAT	893
QY	281	GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaIleArgAsnThr	300
DB	894	GGAAACCCCTCAGAACACTTTGTATCCAGGGCTGATTGATCTTGCCATGCGCAGAACCT	953
QY	301	LeuIleGlyIleAspAspPheIleSerGlyGlnThrIleTyrMetIleSerValIleValAsp	320
DB	954	TTGATTGGTGGAGATGACTTCAAGAGTGTCTCAGACCAAAATGAATCTGTGTGGTGAAT	1013
QY	321	PheLeuValGlyIleGlyIleIleAsnProThrSerIleValSerTyrAsnIleLeuGlyAsn	340
DB	1014	TTCTTTGGGGGGCTGGATATCAAGCAACATCTATAGTCAGTTACACACATCTGGGAAAC	1073
QY	341	AsnAspGlyMetAsnLeuSerIleProGlnThrPheIleArgSerIleGluIleSerIleSer	360
DB	1074	AATGATGGTATGAATCTTGGGCTCCACAACTTTCCTTCCAAAGAAATCTCCAAAGAC	1133
QY	361	AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyIleHisPro	380
DB	1134	AACGTTGTGATGATATGATGTCACAGCATGSCATCTTATGAGCCTGGTGAACATCA	1193
QY	381	AspHisValIleValIleIleTyrgluProTyrValGlyAspSerIleArgAlaMetAsp	400
DB	1194	GACCATGTTGTTATTATTAATGATGATGCTTACGTAGGGGACAGACAGACCATGAT	1253
QY	401	GluTyrThrSerGluIlePheMetGlyGlyIleSerThrIleValLeuHisAsnThrCys	420
DB	1254	GAGTACACTTCAGAGATATTCATGGGTGGAAAGACACCATGTTTGGCAACAACATGC	1313
QY	421	GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleLeuAlaIleLeuSer	440
DB	1314	GAGGATTCCTCTTATAGCTGCTCTTATATCTTGGACTTGGCTCTTCTGCTACGCTCAGC	1373
QY	441	ThrArgIleGluPheIleValGluAsnGluGlyIlePheHisSerPheHisIleProValAla	460
DB	1374	ACTAGATATGAGTTTAAAGCTGAATATGAGGGAATATTCATCTCAACCCAGTTGCT	1433
QY	461	ThrIleLeuSerTyrLeuThrIleValAlaProLeuValProProGlyThrProValIleAsn	480
DB	1434	ACCATCTCTCAGCTTACTCACCAGAGCTCCTCTGTGTCCACCGGATACCACTATGTGTAAT	1493
QY	481	AlaLeuSerIleGluArgAlaMetLeuGluAsnIleMetArgAlaCyValAlaIleLeuAla	500
DB	1494	GCAATGTCAAGCAGCGTGCATGCTGGAAAAATATGAGGGCTTGTGTGATTGGGCT	1553
QY	501	ProGluAsnAsnMetIleLeuGluTyrIleIle	510
DB	1554	CCAGGAATTAATCATGATTTCTGAGATCAAG 1583	

RESULT 5

AAV62440 standard; cDNA; 1782 BP.

AAV62440;

17-OCT-2003 (revised)
 02-FEB-1999 (first entry)

XX Soybean wild-type myo-inositol 1-phosphate synthase cDNA.
 DE Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
 KW phytic acid; db.
 XX
 OS Glycine max; line LR13.
 XX
 FH Key Location/Qualifiers
 FT CDS 54..1586
 FT /tag= a
 PN MO9845448-A1.
 XX
 XX 15-OCT-1998.
 PD
 PF 07-APR-1998; 98MO-US006822.
 XX
 PR 08-APR-1997; 97US-00835751.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PI Hitz WD, Sebaecian SA;
 XX
 XX WPI; 1998-568353/48.
 DR P-PSDB; AAW9740.
 XX
 PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
 PT for generating plants with altered levels of e.g. raffinose, stachyose,
 PT phytic acid, etc.
 PS Example 5; Page 44-45; 63pp; English.
 XX
 CC This is the nucleotide sequence of cDNA encoding the wild-type soybean
 CC myo-inositol 1-phosphate synthase (MI 1-PS) present in clone p5bm1-1ps
 CC (ATCC 97970). The clone was isolated from a cDNA library of soybean line
 CC LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
 CC thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,
 CC raffinose and stachyose. A mutant MI 1-PS nucleic acid (see AAW62443) has
 CC been identified in soybean line LR13, a mutagenised line of low raffinose
 CC saccharide phenotype. Sequencing revealed a single base change mutation
 CC (G to T at base 1241) in the LR13 sequence. The mutation results in a
 CC seed phenotype of very low raffinose saccharide sugars, very high sucrose
 CC and low phytic acid. The nucleic acid is used to alter the raffinose
 CC saccharide, sucrose, phytic acid and inorganic phosphate content of
 CC soybean seeds, leading to useful soybean products, e.g. a seed phytic
 CC acid content of less than 17 ug/g, a seed content of raffinose and
 CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
 CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX Sequence 1782 BP; 516 A; 371 C; 411 G; 484 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 2, 44e-253 Length: 1782
 Score: 2632.00 Matches: 510
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-10-718-952-2 (1-510) x AAW62440 (1-1782)
 QY 1 MetPheIleGluAsnPhelysValIGuCyProAsnVallyrThrGluThrGluIle 20
 DB 54 ATGTTTCATCGAGAAATTTTAAGGTGTGAGTGTCTTAAGTGAAGTCAACCGAGCTGAGATT 113
 QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
 DB 114 CAGTCCTGGTACACTACGAAACACCGAACTTGTTCAAGAAACGGAATGACCTTAT 173
 QY 41 GlnTrpIleValIysProIysSerValIysTyrGluPhelysThrAsnIleHisValPro 60
 DB 174 CAGTGATTTGTCAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCT 233

QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
 DB 234 AAAATTAGGGGTATAGTGTGTGGGTGGGTAAGAAACAGGCTCAACCTTCACGGGTGGT 293
 QY 81 ValIleAlaAsnArgGluGlyIleSerTPrAlaThrLysAspLysIleGlnIleAsn 100
 DB 294 GTTATTGCTTACCGAAGGGCATTTTCATGGGCTTCAAAAGACAAAGTTCAACAGCCAT 353
 QY 101 TyrPheGlySerLeuThrGlnIleSerAlaIleArgValGlySerPheGlnGlyGlu 120
 DB 354 TACTTGCTCCCTCCACCAAGCTCAGCTATCCGAGTTGGGCTCTTCCAGGAGAGAA 413
 QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 DB 414 ATCTATAGCCCATTCYCAAGAGCCTGTCTTCAATGGTTAACTTCAGCAACATTTGTTTGG 473
 QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 DB 474 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCAAGGCAAGGTTTGAC 533
 QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
 DB 534 ATCATTTTCAGAAAGCAGTTGAGGCTTTCATGTGAATCATGCTTCACCTCCCGGAATC 593
 QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
 DB 594 TATGACCCCGAATTTCACTTGCTGCCAACAAAGAGCGGTGCCAAACACCTCATCAAGGGC 653
 QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
 DB 654 ACAAAGCAAGAGCAAGTTCACAAATCATCAAAAGCATCAAGGCGTTTAAAGAGCCACC 713
 QY 221 LysValAspLysValValValIleuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
 DB 714 AAAGTGACAAAGGTGGTGTATCTGTGACTGCACACAGAGAGTACAGTAATTTGGTT 773
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluValGlu 260
 DB 774 GGGGGCTTAAAGACCACTGAGAAATCTTGGCTGCGTGTGACGAATGAGGCTGAG 833
 QY 261 IleSerProSerThrLeuTyrAlaIleAlaCyValMetGluAsnValProPheIleAsn 280
 DB 834 ATTTCTCTTCCACTTGTATGCTGATTCCTTGTATGAAATGTTCTTTCATTAAT 893
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 DB 894 GGAAGCCCTCAGAACACTTTGTACAGAGGCTGATTGATCTTGCCATGCGAGAACACT 953
 QY 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
 DB 954 TTGATTGGTGGAGATGACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTGGTTGAT 1013
 QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
 DB 1014 TTCCTTTGGGGGCTGGTATCAAGGCCACATCTTATGTAAGTTCAACCATCTGGGAAC 1073
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
 DB 1074 AATGATGATGAATCTTCCGCTCCACAAACTTTCCTCCCAAGGAATCTCAAGAGAC 1133
 QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
 DB 1134 AACGTTGTGAAGATATGTGTCAACGACATGCAATCTCTTAAGAGCTGGTGAACATCA 1193
 QY 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
 DB 1194 GACCATTTGTTGTTATTAAGTATGTGCTTACGTAAGGGAACGCAAGAGCCATGAT 1253
 QY 401 GlnTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
 DB 1254 GAGTACACTTCAGAGATATTCATGTGGTGGAAAGACACCATTTGTCACAAACATGTC 1313

QY 421 GUAASPserLeuValaAlaProIleIleLeuAspLeuValleuValaIleuSer 440
DB 1314 GAGGATTCCTCTTACTCTCTCTATTAATCTTGAGCTTGCTCTTGCTGAGCTCAAC 1373
QY 441 THRATGILLeuPheLysAlaGluValGluGlyLysPheHisSerPheHisProValAla 460
DB 1374 ACTAGATATGAGATTAAAGCTGAAATGAGGGGAAATTCACACATTCACACCGAGTTGCT 1433
QY 461 THRileuSerTyrLeuThrLysValaProLeuValProProGlyThrProValValaAsn 480
DB 1434 ACCATCTAGACTACTCTACCAAGGCTCTCTGTTCTCCACGGGTACACCAAGTGGTGAAT 1493
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValaGlyLeuAla 500
DB 1494 GCATTGTCAAGCAGCGCTGCAATGCTGGAATAATATGAGGGCTGTGTGGATTGGCC 1553
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1554 CCAGAGAAATACATGATCTCGAGTCAAG 1583
RESULT 6
AAV62443
ID AAV62443 standard; cDNA; 1533 BP.
XX AAV62443;
AC AAV62443;
XX 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX Soybean mutant myo-inositol 1-phosphate synthase cDNA.
XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid; de.
XX Glycine max; line LR33.
OS WO9845448-A1.
PN 15-OCT-1998.
PD 07-APR-1998; 98WO-US006822.
XX 08-APR-1997; 97US-00835751.
PR (DUPO) DU PONT DE NEMOURS & CO E I.
PA Hitz WD, Sebastian SA;
PI WPI; 1998-568353/48.
XX P-PSDB; AAW79741.
DR Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
XX phytic acid, etc.
PS Example 5; Page 48-49; 63p; English.
XX This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
CC inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
CC clone was isolated from a cDNA library of soybean line LR33 by PCR
CC amplification (see AAV62441-42). Line LR33 was obtained by chemical
CC mutagenesis of wild-type soybean genome and as a reduced raffinose
CC saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. Sequencing revealed a single base change
CC mutation (G to T at base 1241) in the LR33 sequence when compared to the
CC wild-type sequence (see AAV62440). The mutation results in a seed
CC phenotype of very low raffinose saccharide sugars, very high sucrose and
CC low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
CC raffinose saccharide, sucrose, phytic acid and inorganic phosphate
CC content of soybean seeds, leading to useful soybean products, e.g. a seed
CC phytic acid content of less than 17 ug/g, a seed content of raffinose and
CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6,266-253 Length: 1533
Score: 2627.00 Matches: 509
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 2 Gaps: 0
US-10-718-952-2 (1-510) x AAV62443 (1-1533)
QY 1 MetPheIleGluAsnPheLysValaGlyCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTTTCATGAGAAATTTTAAAGGTGAGTGTCTTAATGTGAATGACACCGAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCCGTGTACACTACGAAACCAACGAACTGTTCACGAGAACGGAATGGCACTTAT 120
QY 41 GlnTrrIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATTGTCAAAACCAAAATCTGTCAATACGAATTTAAACCAACATTCATGTTCT 180
QY 61 LysLeuGlyValMetLeuValGlyTrrGlyLysAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTAGGGGTAAATGCTTGTTGGGTGGGTGAAACAAACGGCTCAACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyLysSerTrrAlaThrLysAspLysIleGlnIleAlaAsn 100
DB 241 GTTATGTCTTAACCGAGAGGGCATTTTCATAGGGCTTACAAAGCAAAATTCACAAACCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleAsnValGlySerPheGlnGlyGlu 120
DB 301 TACTTGGCTCCTCAACCCAAAGCTTCAAGTATCCAGTTGGGTCTTCCAGGAGAGGAA 360
QY 121 IleTrrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCCCATTCAGAGCGCTGCTTCAATGTTAACCCCTGACAGACATGTGTTGGG 420
QY 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATCAGCAACATGAACTGTGTATGTCATGCGCAGGGCAAAAGGTGTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTGCAAGAGCATTTGAGGCTTACATGAAATTCATGCTTCCACTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TATGACCGGAAATTTCTATGCTGCTCCAAACAGAGGAGCGTGCACAAACGTCATCAAGGGC 600
QY 201 ThrLysGlnGlnValaGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAAGCAAGACCAAGTTCACAAATCATCAAAAGCATCAAGCGCTTTAAGAACACACC 660
QY 221 LysValaAspLysValaValaLeuTrrPrrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGAAGAAAGGTGTGTACTGTGAGCTGCCAACAGAGAGTACGTAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValaAlaValaAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTTAAAGACACCATGAGATCTCTTGCGTGGCTGTGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCCACCTTGATGCAATGCTTGTTATGAAAAGTTCTTTCAATTAAT 840
QY 281 GlySerProGluAsnThrPheValaProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAAGAACATTTGTATCACGGCGTGAATGATCTTGCCATCGCAGGAACACT 900

```

QY 301 LeuileGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAAGCCAAATGAAATCTGTTGGTTGAT 960
QY 321 PheLeuValGlyValaglyLeuLysProThrSerLysLeuValSerLysAsnIleuGlyAsn 340
DB 961 TTCCTTGGGGGCTGGTATCAAGCCAAATCATATAGTCAGTACCAACATCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyLysSerLys 360
DB 1021 AATGATGGATGATGATCTTCGCTCCCAAACTTCCGTTCCAAAGAAATCTCCAAAGAGC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleuLeuTyrGluProGlyGlyLysPro 380
DB 1081 AACGTTGTTGATATATGTCACAGCAATGCCATCTCTATAGAGCTGGTGAACATCCA 1140
QY 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTGTTATTAATGATAGTCCCTTACGTAGGGGACAGCAATAGCCATGGAT 1200
QY 401 GluTyrThrSerGluLysPheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAAGACCACTGTTTGCACACACATGCT 1260
QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuValGluLeuSer 440
DB 1261 GAGGATTTCCCTTACGCTGCTCATATATCTGATGCTGCTCTCTGCTGAGCTCAGC 1320
QY 441 ThrArgIleGluPheLysValGluAsnGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGATGAGGTTTAAAGCTGAAATGAGGAAATTCACATCTCCACCAAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValAlaAsn 480
DB 1381 ACCATCTCAGCTACCTCAACCAAGGCTCTCTGTTCCACCGGTACACCAAGTGGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTTCAAAGCAGGCTGCATGCTGCGAATAATGAGGCTGTGTGGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1501 CCAGAGATATACATGATTCTCGAGTACAG 1530

```

RESULT 7

ADQ14494
ID ADQ14494 standard; cDNA; 1533 BP.

AC ADQ14494;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.

KM Soybean: myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.

OS Glycine max.
XX Synthetic.

XX Key Location/Qualifiers
FT CDS 1..1533

FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #1"
FT replace(1241,G)
FT /*tag= b

US2004128713-A1.

01-JUL-2004.

XX

```

PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hiltz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-533135/51.
DR P-PSDB; ADQ14495.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Claim 10; SEQ ID NO 5; 48bp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.: 6,26e-253 Length: 1533
Score: 2627.00 Matches: 509
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 12 Gaps: 0

US-10-718-952-2 (1-510) x ADQ14494 (1-1533)

```

QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTCATCGAGAAATTTTAAAGTTGAGTCTCTAATGTGAAGTACACCAAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCTGATACACTACCAACCAACCACTGTTCCAGAAACAGAGATGACACTAT 120
QY 41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATTGTCAAACCAACCAATCTGTCAAAATGCAATTTAAACCAACATCCATGTTCT 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AATTTAGGGGTATGCTGCTGCTGGGTGGAAACACGCTCAACCTCACCCTGGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100

```

```

Db      241 GTTATTGCTAACCGAGGCGATTTCATGGGCTACAAAGACAGATTCAACAAACCAT 300
Qy      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db      301 TACTTTGGCTCTCTTACCCCAAGCCTCAGCTATCCAGATTGGGTCTTCCAGGGAAGGAA 360
Qy      121 IleTyrAlaProPheLeuSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db      361 ATCTATGCCCATTCAGAGAGCGCTGCTTCAATGGTTAACTTGACGACATGTGTGGGG 420
Qy      141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db      421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCTCATGCGCAGGCGCAAGGTGTTGAC 480
Qy      161 IleAspLeuGlnLeuGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
Db      481 ATCGATTTCAGAGAGCGTTGAGGCTTACATGAAATCCATGCTTCCATCCCGGAAATC 540
Qy      181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
Db      541 TATGACCCGGATTTCATTGCTCTCCAAACAAAGAGCGTGCAACAAAGTCATCAAGGCG 600
Qy      201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
Db      601 ACAAAGCAGAGACAGTTCACAAATCATCAAGACATCAAGCGCTTTAAGAGAGCCACC 660
Qy      221 LysValAspLysValValIleuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      661 AAAGTGGACAGAGGTGTTACTGTGATGCTGCCACACAGAGAGGTACGTAATTTGGTT 720
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
Db      721 GTGGGCTTAATGACACCATGAGAAATCTTGCTGCTGTGACAGAAATAGAGCGTGAG 780
Qy      261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      781 ATTTCTCTTCCACCTTGATGCAATGCTGTGTATGTAAGTAAAGTTCCTTCAATTAAT 840
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCGTGAATTGATCTTGCCATCCCGAGAACACT 900
Qy      301 IleuIleGlyIAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      901 TTGATTGGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGGTGAAT 960
Qy      321 PheLeuValGlyValAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db      961 TTCCTTGTGGGGCTGTGTCAAGCCAAACATCATATAGTACAAACCATCTGGGAAAC 1020
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db      1021 AATGATGTAATGATTTTCGGCTCCAAACCTTTCCTTCCAAAGAAATCTCCAAAGAGC 1080
Qy      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db      1081 AACGTTGTGATGATGTGTCAACAGCAATGCCATCTCTATGAGCGCTGTGAAATCTCA 1140
Qy      381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1141 GACCAGTGTGTGTAATTAAGATATGCTTACGTATGGGGAACAGCAATGAGCCATGAGAT 1200
Qy      401 GluTyrThrSerGluLeuPheMetGlyLysSerThrIleValLeuHisAsnThrCys 420
Db      1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGACCAATGTTTGTGCAACACATAC 1260
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1261 GAGGATTTCCCTTACTGCTCTATTATCTTGACATTTGATGCTCTTCTGTGAGCTACAC 1320
Qy      441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      1321 ACTAGATCGAGTTTAAAGCTGAATAATGAGGGAATAATTCACATCTCAACCCAGTTGCT 1380

```

```

Qy      461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
Db      1381 ACCATCTCAGTACTACTCACCAGAGCTCTCTGTTCCACCGGGTACACAGTGTGTAAT 1440
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetAlaArgAlaCysValGlyLeuAla 500
Db      1441 GCATTGTCAAGCAGCGGCAATGCTGGAAAAACATTAATGAGGCTGTGTGGATTTGGCC 1500
Qy      501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db      1501 CCAGGAATTAACATGATTCTCGAGTACAG 1530

RESULT 8
ADS81997
ID      ADS81997 standard; cDNA; 1533 BP.
XX
AC      ADS81997;
XX
DE      18-NOV-2004 (first entry)
XX
KW      Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.
XX
XX      Glycine max; line LR33.
XX      Synthetic.
OS
XX
FT      Key Location/Qualifiers
FT      CDS 1..1533
FT      /*tag= a
FT      /product= "myo-inositol 1-phosphate synthase"
FT      replace(1188,G)
FT      mutation
FT      /*tag= b
XX
PN      US2003074685-A1.
XX
PD      17-APR-2003.
XX
PF      11-MAR-2002; 2002US-00025003.
XX
PR      08-APR-1997; 97US-00835751.
XX      07-APR-1998; 98WO-US006822.
XX
PA      (HITZ/) HITZ W D.
PA      (SEBA/) SEBASTIAN S A.
XX
PI      Hitz WD, Sebastian SA;
XX
XX      MPI; 2004-639957/62.
XX      P-PSDB; ADS81998.
XX
PT      Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT      phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT      phytic acid and inorganic phosphate content of soybean seeds.
XX
PS      Claim 8; SEQ ID NO 5; 34pp; English.
XX
XX      The invention relates to an isolated nucleic acid fragment encoding a
XX      soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX      phosphate synthase having decreasing capacity for the synthesis for myo-
XX      inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX      nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX      complement, subfragment or the complement of the subfragment, operably
XX      linked to suitable regulatory sequences, where expression of the chimeric
XX      gene results in a decrease in expression of an endogenous or native gene
XX      encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX      comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX      acid content of less than 17 micromol/g, a seed content of raffinose plus
XX      stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX      greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX      from the plant, making a soybean plant with the heritable phenotype

```

CC (comprising crossing LR3 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with a
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence encodes a mutant myo-
 CC inositol 1-phosphate synthase.

XX Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6,266-253	Length:	1533
Score:	2627.00	Matches:	509
Percent Similarity:	99.80%	Conservative:	0
Best Local Similarity:	99.80%	Mismatches:	1
Query Match:	13	Indels:	0
DB:	13	Gaps:	0

US-10-718-952-2 (1-510) x ADS81997 (1-1533)

QY 1 MecPhe11eGluAAspPheVal1GluCyPProAsnVal1SerThrGluThrGlu1le 20
 DB 1 AGTTTATGAGAAATTTTAAGTTGAGTGTCTTAATGTAAATGACCGAGATCGAGATT 60
 QY 21 GlnSerVal1TyraSerThrGluThrGluLeuVal1HisGluAAsnArgAsnGlyThrTy 40
 DB 61 CAGTCCGGTGTACAACTACGAAACACACCACTGTTCACGAGAAAGAAATGCGACTAT 120
 QY 41 GlnThr11eVal1LeuPhePro1LeuVal1SerThrGluPheThrAsn1Leu1Pro 60
 DB 121 CAGTGGATGTCAAAACCAATCTGTCAAAATGAAATTTTAAACCAACATCCATGTTCT 180
 QY 61 LysLeuGlyVal1MetLeuVal1GlyTrpGlyGlyAsnAsnGlySerThrLeuThrGly 80
 DB 181 AAATTAGGGGTAATGCTTGTGGGTGGGATGAAACACCGCTCAACCTCAGCGGTGCT 240
 QY 81 Val11eAlaAsnArgGluGly11eSerThrPalaThrLysAspLys11eGln1AlaAsn 100
 DB 241 GTTATTGTACACGAGAGGCGCATTTGATGGGCTAACAAAGGACAAAGATTCACAAAGCCAA 300
 QY 101 TyrPheGlySerLeuThrGln1AlaSer1Ala1eArgVal1GlySerPheGlnGlyGlu 120
 DB 301 TACTTGGCTCCTCCACCAAGCTCAGCTATCCGAGTGGGCTCTTCCAGGGAGAGAA 360
 QY 121 IleTyra1ProPheLysSerLeuLeuProMetVal1AsnProAspAsp11eVal1PheGly 140
 DB 361 ACTTATGCCCATCTCAAGAGCCCTGCTTCCAATGGTTAAACCTGACACATGTGTTGGG 420
 QY 141 GlyTrpAsp11eSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysVal1PheAsp 160
 DB 421 GGATGGGATATCAGCAATGAACTGAGCTGATGCGCATGGCCAGGCAAAAGGTTTGAC 480
 QY 161 IleAspLeuGln1LeuGln1LeuArgPro1TyrMetGluSerMetLeuProLeuProGly11e 180
 DB 481 ATCGATTTCAGAAAGAGTTGAGGCTTCAATGAAATCCATGTTCCACTCCCGGAATC 540
 QY 181 TyrAspProAspPhe11eAla1AlaAsnGlnGluGluArgAlaAsnAsnVal11eLysGly 200
 DB 541 TATGACCCCGATTTCATTCTGCTGCCAACAGAGAGACGTCGCAACACGTCATCAAGGGC 600
 QY 201 ThrLysGlnGluGlnVal1Gln11e11eLysAsp11eLysAlaPheLysGlu1AlaThr 220
 DB 601 ACNAAAGCAAGACAGTTCAACAAATCATCAAAAGACATCAAGGCGTTTAAAGNAGCCACC 660

QY 221 LysVal1AspLysVal1Val1LeuThr1AlaAsnThrGluArgTyrSerAsnLeuVal 240
 DB 661 AAGTGGACAAAGGTGGTGTACTGTGACTGCTCCAAACAGAGAGGTACAGTAATTTGGTT 720
 QY 241 Val1GlyLeuAsnAspThrMetGluAsnLeuAla1Val1AspArgAsnGlu1Gly 260
 DB 721 GGGGCTTAAATGACACCATGAGAAATCTCTTGCTGCTGTGGACAGAAATGAGCTGAG 780
 QY 261 11eSerProSerThrLeuThrAla11eAlaCyVal1MetGluAsnVal1ProPhe11eAsn 280
 DB 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGTATGAAATGTTCTTTTCATTAAAT 840
 QY 281 GlySerProGluAsnThrPheVal1ProGlyLeu11eAspLeuAla11eAlaArgAsnThr 300
 DB 841 GGAAGCCCTCAGAAACCTTTGTACAGGGCTGATTTGATTCCTTCATCCGAGAAACACT 900
 QY 301 Leu11eGly1AspAspPheLysSerGlyGlnThrLysMetLysSerVal1LeuVal1Asp 320
 DB 901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTGTTGGTAT 960
 QY 321 PheLeuVal1GlyAlaGly11eLysProThrSer11eVal1SerTyrAsnHisLeuGlyAsn 340
 DB 961 TTCCTTGTGGGGCTGGTATCAAGCCCAACATCTATAGTCAGTTACAAACCATCTGGGAAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlu11eSerLysSer 360
 DB 1021 AATGATGGTATGAATCTTGGCTCCAAACATTTCCGTTCCAAAGAAATCTCCAGAGC 1080
 QY 361 AsnVal1Val1AspAspMetVal1AsnSerAsnAla11eLeuTyrGluProGlyGluHisPro 380
 DB 1081 AAGCTTGTGATGATATGTGTCACAGCATGCGCATCTCTTAAGAGCTGTGTAACATCCA 1140
 QY 381 AspHisVal1Val11eLysTyrVal1ProTyrVal1GlyAspSerLysArgAlaMetAsp 400
 DB 1141 GACCATGTTGTTTATTAATGATGTCCTTACGAGGAGACAGCAATGAGCCATGAGAT 1200
 QY 401 GlyTyrThrSerGln1PheMetGlyGlyLysSerThr11eVal1LeuHisAsnThrCys 420
 DB 1201 GAGTACACTTCAGAGATATTCAGGTGGAAGAGACACATGTTTGGCAACACATGC 1260
 QY 421 GluAspSerLeuLeuAla1Pro11e11eLeuAspLeuVal1LeuLeu11eGluLeuSer 440
 DB 1261 GAGGATCCCTCTTAGCTCTCTATTAATCTTGAGCTTGCTCTTCTCTAGCTCAGC 1320
 QY 441 ThrArg11eGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProVal1Ala 460
 DB 1321 ACTAGATATGAGTTTAAAGCTGAAATGAGGAAATTCACCTCATTCACCCAGATTGCT 1380
 QY 461 Thr11eLeuSerTyrLeuThrLysAlaProLeuVal1ProProGly1ThrProVal1AlaAsn 480
 DB 1381 ACCATCTCAGCTACCTCACCAAGGCTCTCTGTTCCACCGGGTACACCAAGTGTGAAT 1440
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsn11eMetArgAlaCyVal1GlyLeuAla 500
 DB 1441 GCATTGTCAAACACAGCGTGCATGCTCGAAACATATGAGGCTTGTGTGGATTGGCC 1500
 QY 501 ProGluAsnAsnMet11eLeuGluTyrLys 510
 DB 1501 CCAGAGAAATTAATGATATCTCGAGTACAAAG 1530

RESULT 9

ADQ14504
 ID ADQ14504 standard; cDNA; 1533 BP.

AC ADQ14504;

XX 23-SBP-2004 (first entry)

DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.

XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;

KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 inorganic phosphate.

DB 1141 GACCATGTTGTTATTAAGTAGTGCTTACAGGAGGATGACAGAGCCATGGAT 1200
QY 401 GUTYThrSerGluIlePheMetGlyGlyLysSerThrIleValIleuHisLeuThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGGAAAGAACCCATGTTGTTGGCAACACATGT 1260
QY 421 GUAASPLeuLeuLeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTTAGCTGCTCTTATTTATCTGACATTTGTTGCTTCTGCTGAGCTGAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGATTCAGATTAAAGCTGAAGGAAATGAGGAAATTCACATCATTCACCCAGTTGCT 1380
QY 461 ThrIleLeuSerTyrIleuThrLysAlaProLeuValProProGlyThrProValIleuAsn 480
DB 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGTTCCACGGGCTTACACAGTGGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleuMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTCTCAAGACGCTGCAATGCTGGAAACATTAATGAGGCTTGTGGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1501 CCAGAGATTAACATGATTCTCGAGTCAAG 1530
RESULT 10
ADQ14502
ID ADQ14502 standard; cDNA; 1533 BP.
XX
AC ADQ14502;
XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
XX
KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
OS Glycine max.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #4"
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 16-APR-1999; 99US-0029315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
WI MPI, 2004-533135/51.
XX
DR P-PSDB; ADQ14503.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean

PT products.
XX
PS Example 8; SEQ ID NO 13; 48bp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native plant
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant,
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.17e-250 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservatve: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
DB: 12 Gaps: 0

US-10-718-952-2 (1-510) x ADQ14502 (1-1533)
QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 AGTTATATGAGAGATTTTAAGGAGAGAGTCTTAATGTAAGTACCCAGATGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCCGTGTACACACTAGCAAGAAACCGAATCTTGTTCAAGAGACAGATGGCACTAT 120
QY 41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGAATGTCAAACCCAAATCCGTCACATCTTAATTAACCAACCCATGTTCCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyLysAsnAsnGlySerThrLeuThrGlyIle 80
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGAGGAAACAGCGCTTACCTCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTAACAGAGAGGCACTTTCATGGGCTACAAAGACAAGTTCAACAAGCCAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTCACCAAGCCTCAGCTATTCGAGTTGATTCCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCCCATTCAGAGCTGCTTCCAAAGGTTAATCTGACACACATGTGTTGGG 420
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGAGATACGACACATGAACCTGCTGATGCAATGCCAGGCAAGAGGTGTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTCAGAGAGAGTGAAGGCTTACATGGAATCATGATGTTCCATCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAsnGlnGluLysValGlnAlaAsnValIleLysGly 200


```

Db      541 TACGACCCGGATTTCATTCCTGCCAACAGAGAGCGTCCAGAACAGTCGATTAAAGGC 600
Qy      201 ThrValGlnGluGlnValGlnGlnIleIleValAspIleIleValAlaPheValGlnAlaThr 220
Db      601 ACAAGACAGACAGACAGATTCAGCAATCATCAAGACATCAAGACCGTTTAAAGAACCCACC 660
Qy      221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      661 AAAGTGGACAAAGTGGTGTCTCTGTGACCTGCACAACAGAGAGATATGCAATTTGGTT 720
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
Db      721 GTAGGCTTAAATGACACCACTGAGAAATCTCTGGCTGCTGTGACAGAAATATAGGCTGAG 780
Qy      261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      781 ATTTCTCTCCATCCATCTGTATGCACTTCCCTGTGATGAAAGATGTTCTTTCATTAAAT 840
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      841 GGAGACCCCTCAGAACACCTTTGTGACAGGGCTGATTGATCTTGCCATCGCAGAGAACT 900
Qy      301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      901 TTGATTGGTGAAGATGACTTCAAGAGTGGTCAAGACCAAAATGAAATCTGTGTGGTGTAT 960
Qy      321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleGluAlaAsn 340
Db      961 TTTCTTTGGGGGCTGGTATCAAGCCACATCTATAGTTACACCATCTGGGAAC 1020
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db      1021 AATGATGATGATATCTCGGCTCCACAACTTCGCTCCAGAGAAATCTCCAAAGAC 1080
Qy      361 AsnValValAspAspMetValAsnSerAlaAlaIleLysTyrGluProGlyGluHisPro 380
Db      1081 AACGTTGTGAGATATGCTCAACAGCAATGCCATCTTATGAGCGTGTGAAACATCCC 1140
Qy      381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1141 GACCATGTTGTTTATTAAGATATGCTTACGTAGGGGATAGCAGAGAGCCATGAT 1200
Qy      401 GlyTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db      1201 GAGTACACTTCAAGATATTCATGCGTGAAGAAACACCATTTGTTGCACAAACATGT 1260
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1261 GAGGATTCCTTTTACCTCTCTATATCTTGAAGCTTGTCTTCTGCTAGAGCTGAGC 1320
Qy      441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      1321 ACTAGATTCAGTTTAACTGAAAATGAGGAAAATTCACATCATCCACCACTGAGTGT 1380
Qy      461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValAlaAsn 480
Db      1381 ACCATCTCAGCTATCTGACCAAGGCTCTCTGCTGCCACCGGATCACCACTGGTGAT 1440
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1441 GGAATGTCAACAGCGTGCATGCGAAGAAACATATATAGAGGCTTGTGTGATTTGGCC 1500
Qy      501 ProGluAsnAspMetIleLeuGluTyrLys 510
Db      1501 CCGAGAAATTAACATGATCTCTCGAGTACAG 1530

```

RESULT 11

ADS82005 ID ADS82005 standard; cDNA; 1533 BP.

XX AC ADS82005;

XX DT 18-NOV-2004 (first entry)

```

XX DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
XX KW Soybean; plant; myo-inositol 1-phosphate synthase; ssr gene; phytic acid;
XX KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX OS Glycine max; cultivar Wye.
XX FT Key Location/Qualifiers
XX FT CDS 1..1533
XX FT /tag=a
XX FT /product="myo-inositol 1-phosphate synthase"
XX PN US2003074685-A1.
XX PD 17-APR-2003.
XX PF 11-MAR-2002; 2002US-00025003.
XX PR 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX PA (HITZ/) HITZ W D.
XX PA (SEBA/) SEBASTIAN S A.
XX PI Hitz WD, Sebastian SA;
XX DR WPI: 2004-639957/62.
XX DR P-PsDB; ADS82006.
XX PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX PT phytic acid and inorganic phosphate content of soybean seeds.
XX PS Claim 2; SEQ ID NO 15; 34pp; English.
XX PS
XX CC The invention relates to an isolated nucleic acid fragment encoding a
XX CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX CC phosphate synthase having decreasing capacity for the synthesis for myo-
XX CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX CC complement, subfragment or the complement of the subfragment, operably
XX CC linked to suitable regulatory sequences, where expression of native gene
XX CC gene results in a decrease in expression of an endogenous or native gene
XX CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX CC acid content of less than 17 micromol/g, a seed content of raffinose plus
XX CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX CC from the plant, making a soybean plant with the heritable phenotype
XX CC (comprising crossing LR33 or the plant comprising the chimeric gene with
XX CC an elite soybean plant and selecting a progeny plant of the cross of
XX CC crossing step that has a heritable phenotype as mentioned above), seeds
XX CC of soybean plant made by the above method, a soy protein product derived
XX CC from seeds of a soybean plant (homologous for one or more gene encoding a
XX CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
XX CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
XX CC heritable phenotype as mentioned above), and making or producing a
XX CC soybean protein product derived from seeds of a soybean plant with
XX CC heritable phenotype as mentioned above. The nucleic acid is useful for
XX CC altering raffinose saccharide, sucrose, phytic acid and inorganic
XX CC phosphate content of soybean seeds thus leading to valuable and useful
XX CC soybean products, since the presence of high concentration of raffinose
XX CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
XX CC when consumed by humans. The present sequence encodes a wild-type myo-
XX CC inositol 1-phosphate synthase.
XX SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

```

Alignment Scores:

```

Pred. No.: 3 176-250 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4

```


Best Local Similarity: 98.63% Mismatches: 3
 Query Match: 98.78% Indels: 0
 DB: 13 Gaps: 0
 US-10-718-952-2 (1-510) x ADS82005 (1-1533)

```

QY 1 MetPheIEGLuAnPheLysValIGlCyProAsnValIleYrThrGluThrGluIle 20
DB 1 AGTTTCATGAGAAATTTTAAAGTAGAGCTCTTAATGTGAAGTACCGAGACTGAGATT 60
QY 21 GlnSerValTYrAsnTYrGluThrThrGluLeuValHISGluAnbArgAsnGlyThrTYr 40
DB 61 CAGTCGCTGACAACTACGAAACACCGAAGCTTTCACGAGAAACAGGAATGCACTTAT 120
QY 41 GINTPILLeValLysProLysSerValLysTYrGluPheLysThrAniIleHISValPro 60
DB 121 CAGTGAATGTCAAAACCAATCCGTCACTCAATTTAAACCAACCCCACTGTTCGA 180
QY 61 LysLeuGlyValMetLeuValIGlYTPGlyGlyAsnAsnGlySerThrLeuThrGly 80
DB 181 AAATGGGGGTGATGCTGTGGGTGGGTGAAACACGGCTTACCTCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTYrAlaThrLysAspLysIleGlnIleAlaAn 100
DB 241 GTTATTGCTAACGAGAGGGCAATTCATGGGCTTACAAAGGACAAAGATTCAACAGCCAT 300
QY 101 TYrPheGlySerLeuThrGlnAlaSerAlaIleArgValIGlySerPheGlnGlyGlu 120
DB 301 TACCTTGGCTCCCTCCACCAAGCTCAGCTATTCGAGTTGATCCCTTCAAGGAGAGAA 360
QY 121 ILeTYrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTTCAAGAGCTGCTTCCATGTGTTAACTCGACGACATGTGTTGGG 420
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATCAGCAACATGAACCTGGGTGATGCCATGGCCAGGCAAAAGGTTTGAC 480
QY 161 ILeAspLeuGlnLysGlnLeuArgProTYrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTTCAGAAAGCAGTTGAGGCTTTCATGGAATTCATGTTCCACCTCCCGGAAATC 540
QY 181 TYrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TACGACCCGGAATTTCACTGCTGCAACCAAGAGACGCTGCAACAGCGATTAAGGAGC 600
QY 201 ThrLysGlnGluGlnValGlnIleIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAAAGCAAGACCAATTCAGCAAAATCATCAAAAGCATCAAGCCGTTTAAAGAACCAAC 660
QY 221 LysValAspLysValValLeuTYrThrAlaAsnThrGluArgTYrSerAnLeuVal 240
DB 661 AAAGTGAACAAAGGTGTCTCTGCTGACCTGCCAACACAGAGGATAGCAATTTGGTT 720
QY 241 ValIGlyLeuAnAspThrMetGluAnLeuLeuAlaValAspArgAsnGluAlaGlu 260
DB 721 TAGAGGCTTAAATGACCAATGAGAAATCTTGGCTGCTGAGCAAGAAATGAGGCTGAG 780
QY 261 ILeSerProSerThrLeuTYrAlaIleAlaCysValMetGluAnValProPheIleAn 280
DB 781 ATTTCTCTTCCACCTTGTATGCCATTCCTGTGTATGGAATAATTTCTTTCAATTAAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATGATCTTGCATCCGAGAACTACT 900
QY 301 LeuIleGlyLysAspAspPheLysSerGlyInThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTTGGTGAATGACTTCAAGAGTGTCAAGCAAAATGAATACTGTGTGTGTGAT 960
QY 321 PheLeuValAlaGlyIleLysProThrSerIleValSerTYrAniIleLysGluIleAn 340
DB 961 TTTCTTTGGGGGCTGTATCAAGCCAACTATATAGTTAGTACCAACATCTGGGAAAC 1020

```

```

QY 341 AsnAsnGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGTATGAATCTCTGGCTCCCAAAACCTTCGCTCCCAAGGAAATCTTCCAAAGAC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTYrGluProGlyGluHISPro 380
DB 1081 AACGTTGTGACGATATGTGTCAACAGCATGTCATCTTATGAGCTGTGTGAACATCCC 1140
QY 381 AspHisValValIleLysTYrValProTYrValIGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTGTTGTTATTAATGATGTGCTTACGTAGGGAGATAGCAAGAGCATGAT 1200
QY 401 GlnTYrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHISAnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAACCACTGTTTGGCAACAACATGT 1260
QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTTTTACCTCTCTATTAATCTTGAGACTTGTCTTCTGAGCTGAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHISerPheHISerProValAla 460
DB 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCACATTCACACCAAGTTGCT 1380
QY 461 ThrIleLeuSerTYrLeuThrLysAlaProLeuValProProGlyTYrProValValAn 480
DB 1381 ACCATTTCTAGATATCTGACCAAGGCTCCTGTGTTCCACCGGTACACAGGTGTGAAT 1440
QY 481 AlaLeuSerLysGlyArgAlaMetLeuGluAniLeuMetArgAlaCysValIGlyLeuAla 500
DB 1441 GCATTTCTCAAAACAGCGTCAATCTGGAAGAAACATATAGAGGCTTGTGTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTYrLys 510
DB 1501 CCAGAGAAATACATGATTCCTGAGTACAG 1530

```

RESULT 12
 ADS82003
 ID ADS82003 standard; cDNA; 1533 BP.
 XX
 AC ADS82003;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
 XX
 KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 XX
 OS Glycine max; line 290180P03.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 XX
 PN US2003074685-A1.
 PD 17-APR-2003.
 XX
 PF 11-MAR-2002; 2002US-00025003.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PI Hitz WD, Sebastian SA;
 XX
 DR WPI; 2004-639957/62.
 DR P-PsDB; ADS82004.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX Example 8, SEQ ID NO 13, 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 3,17e-250 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
DB: 13 Gaps: 0

US-10-718-952-2 (1-510) x ADS82003 (1-1533)

QY 1 MetPheIIeGluAsnPhelysValGluCySProAsnValIyetrThrgluThrgluile 20
DB 1 AAGTTATGAGAAATTTTAAAGTAGAGAGTCCTAATGTGAAGTACCGAGCTGGAAT 60
QY 21 GlnSerValTyraSnTyrgluThrgluIeuValHieGluAsnArgAsnglyThryr 40
DB 61 CAGTCCGTGACAACTACGAACCAACGAACTGTTCACAGAAACAGGATGCACTAT 120
QY 41 GlnTPillleValIySProIySerValIyTyrgluPhelysThrsnIleHsValPro 60
DB 121 CAGTGAATGTCAAAACCAATCCGTCAACTAAATTAAACCAACCCATGTTTCCA 180
QY 61 LysLeuGluValMetIeuValGluTyrgluIyAsnAsnglySerThrIeuThrgluIy 80
DB 181 AAATTGGGGGTGAGCTTGTGGGTGGGTGGAACACGAGCTTACCTCCACCGGTG 240
QY 81 ValIleAlaAsnArgIyGluIyIleSerTPAlaThrLysAspIyIleGlnIleAlaAsn 100
DB 241 GTTATTGCTAAACAGAGAGGCAATTCATGGCTACAAAGACAAGATTCAACAGCCAAT 300
QY 101 TyrPheGlySerIeuThrgluIleAsnAlaIleArgValGlySerPheGlnIyGluIu 120
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTTCCAGGAGAGGAA 360

QY 121 IleTyraIaProPhelySerIeuIeuProMetValaAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTCACAAAGTCTGCTTCCAAATGTTAATCCGACGACATGTTGTGGG 420
QY 141 GlyTPAspIleSerAsnMetAsnIeuAlaAspAlaMetAlaArgAlaIlyValPheAsp 160
DB 421 GGATGGGATATCAGAAACATGAACTGGCTGATGCAAGGCGCAAAAGGTGTTGAC 480
QY 161 IleAspLeuGlnIyGluIeuArgProTyrmecGluSerMetIeuProIeuProGlyIle 180
DB 481 ATCGATTTCAGAACAGCTTGAGGCTTACATGAAATCATGTTCCACTCCCGGAGATC 540
QY 181 TyraSPProAspPheIleAlaIleAsnglnIyGluArgAlaAsnAsnValIleIyGly 200
DB 541 TACGACCCGGATTTATTGCTGCTCCAAACCAAGAGAGCGTCCAAACAGTGAATTAAGG 600
QY 201 ThrIyGlnIyGluIyGlnIleIleIleIyAspIleIyValaPheIyGluIaThr 220
DB 601 ACAAAGCAAGAGCAAGTTCAGCAAAATCATCAAAAGCATCAAGCGTTTAAAGAAAGCCACC 660
QY 221 LysValaAspIySValIleuTrpThraIleAsnThrgluArgTySerAsnIeuVal 240
DB 661 AAAGTGCACAAAGTGTGTCTGTGGAGTCCCAACAGAGGTTATGCAATTTGGTT 720
QY 241 ValGlyIeuAsnAspThrMetGluAsnIeuIleAlaIleAspArgAsngluIleGlu 260
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTCTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrIeuTyraIleAlaIleCySValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCCACCTTGATGATGCCATCTGTGTGAAGAAAGTTCCTTTCATTAAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyIleuIleAspIleuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAAACCTTTTGTACAGGGCTGATGATCTTGCCATCGCAGAGAACT 900
QY 301 LeuIleGlyIyAspAspPheIySergIyGlnThryIyMetIySergValIeuValaAsp 320
DB 901 TTGATTTGGTGGAGATGACCTTCAAGAGTGTGACCAAAATGAATCTGTGTGTTGAT 960
QY 321 PheIeuValGlyAlaGlyIleIySProThrsnIleValSerTyraSnHieIeuIyAsn 340
DB 1081 AACGTTGTTGACGATATGCTCAACAGCAATGCAATCTCTATGAGAGCTGTGTAACATCCC 1140
QY 381 AspHieValIleValIleIySryTyraProTyraIyIyAspSerIyAspArgAlaMetAsp 400
DB 1141 GACCAATGTTGTATTAAAGATATGCTTACGTAAGGGAATGAGAAAGACCCATGTGAT 1200
QY 401 GlyTyrrThrsSergIlePheMetGlyIyIySerThrIleValIleuHieAsnThryCyS 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAAACACATTTGTTTGCAACACATGT 1260
QY 421 GluAspSerIeuIleuAlaIleProIleIleuAspIeuValIleuIleuIleuSer 440
DB 1261 GAGGAATCCCTTTTATGCTGCTCTATTAATCTTGAGACTTGCTCTTGCTGAGCTGAGC 1320
QY 441 ThrAArgIleGluPhelysAlaGluAsngluIyIySPhenIleSerPheHieProValAla 460
DB 1321 ACTAGAAATCCAGTTTAAAGCTGAATAATGAGGAAAAATTCATCTCAATCCCAAGTTGCT 1380
QY 461 ThrIleIeuSerTyrrIeuThrIySAlaProIeuValaProProGlyTyrrProValIleAsn 480
DB 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGTGTTCCACGGGATACACCAATGTTGAT 1440

QY 481 AAlaSerIyGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTTCGAAGCAGCGCTGCAATGCTGGAAAACATATAGGCGCTGTGGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTyrIyG 510
DB 1501 CCAGAGATATACATGATTCCTCGAGTACAG 1530

RESULT 13
ADQ14500
ID ADQ14500 standard; cDNA; 1533 BP.
XX
XX ADQ14500;
XX
XX 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
XX Glycine max.
XX
XX Synthetic.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX /tag= a
XX /product= "Mutant soybean myo-inositol 1-phosphate
XX synthase #3"
XX
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX PR 26-APR-1998; 98MO-US006822.
XX PR 26-APR-1999; 99US-00299315.
XX PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG,
XX
XX MPI: 2004-53335/51.
XX P-PSDB; ADQ14501.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Claim 10; SEQ ID NO 11; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with

CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.59e-249 Length: 1533
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.52% Indels: 0
DB: 12 Gaps: 0
US-10-718-952-2 (1-510) x ADQ14500 (1-1533)
QY 1 MetPheIleGluAsnPheIyValGluCysProAsnValIySerThrGluThrGluIle 20
DB 1 ATGTTTCATGAGAAATTTTAAAGTAGAGACTCTTAATGTAAAGTACACCGAGCTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrgIuLeuValIleGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCTTACACATACGAAACCAACCGAATCTTGTTCACGAGAACAGATGGCACTAT 120
QY 41 GlnTPRILEVALIYEPProIySerValIySyrGluPhuIyThrAsnIleHsValPro 60
DB 121 CAGTGAATGTCAAACCCCAATCCGTCACATCTTAAACCAACCCCAATGTTTCCA 180
QY 61 LybLeuGlyValMetLeuValGlyTyrGlyIyGluAsnAsnGlySerThrLeuThrGly 80
DB 181 AAATTGGGGGTGATGCTTGTGGGTTGGGGTGGAACACACGCTTACCTCCACGGGTGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTPRAlaThrIyAspIyValIleGlnIleAlaAsn 100
DB 241 GTTATTGCTAACAGAGAGACATTTTCATGCGCTACAAAGACAAAGATTCACAAAGCCAA 300
QY 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTCACCCCAACCTCAGCTATTCGATTGATCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaPProPheIySerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAAGGTTAATCTCGACACATTTGTGTTGG 420
QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIyValPheAsp 160
DB 421 GGATGGATATACAGCAATGAACCTGCTGATGTCATGGCCAGGCAAAAGGTTTGAC 480
QY 161 IleAspLeuGlnIyGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTCAGAAAGAGATTTGAGGCTTACATGGAATCATGATCTCCCGCAATC 540
QY 181 TyrAspProAspPheIleAlaIleAsnGlnIyGluArgAlaAsnAsnValIleIyGly 200
DB 541 TACGACCCGGAATTCATGCTGCTCCAAACAGAGAGCGGCCCAACACATGATTAAGGGC 600
QY 201 ThrIyGlnIyGlnValGlnIleIleIySAspIleIySAlaPheIySgluIleArh 220
DB 601 ACAAAGCAAGACCAAGATTGACCAATCATCAAGACATCAAGGCGTTTAAAGAGCACCC 660
QY 221 LybValAspIyValIyValIleuTPRAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGACAAAGGTGTTCTCTGGGACTGCAACACAGAGAGATGCAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValIleValIyAspArgAsnGluIle 260
DB 721 GTAGGCTTAAAGACACATGAGAAATCTTTGGCTGCTGTGACAGAAATAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCCTTCACCTTGATGTCATTCGCTGTGTGATGGAAGAAATGTTCTTCAATAT 840

```

Oy 281 GlysSerProGlnAaThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAnthr 300
Db 841 GGAAGCCCTCAGAACCTTTGTACAGGGCTGATGATCTTGCCATCGCAGAACACT 900
Oy 301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTGGGAGAGAGACTTCAAGAGTGGTCAGAACCAAAATATAATCTGTGTGGTGGAT 960
Oy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleSerGlyAsn 340
Db 961 TTCTCTTGCGGGGCTGGTATCAAGCCCACTATATGTTAGTTACACCACTCGGGAAC 1020
Oy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysSer 360
Db 1021 AATGATGATGATGATATCTCGGCTCCACAACTTCCTCCAGAGAAATCTCCAAAGAC 1080
Oy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlyLysPro 380
Db 1081 AACGTTGTTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAACATCC 1140
Oy 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAGGGGATACAGAGAGCCATGAT 1200
Oy 401 GlyTyrThrSerGlyIlePheMetGlyGlyLysSerThrIleValLeuHisAnthrCys 420
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAAGACCATTTGTTGCAACAACATGT 1260
Oy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGlyLeuSer 440
Db 1261 GAGGATTCCTTTAGCTCTCTATTAATCTTGGATTTGGTCTCTTCTCTAGCTGAGC 1320
Oy 441 ThrArgIleGlyPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAGATATCAAGTTAAAGCTAAATGAGGAAATTCACATCTCAACCACTGTTGCT 1380
Oy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATCTCAGCTATCTGACCAAGGCTCCTCTGCTTCCACCGGTACACCAAGCTGGAAT 1440
Oy 481 AlaLeuSerLysGlyArgAlaMetLeuGluAsnIleMetArgAlaGlyValGlyLeuAla 500
Db 1441 GCATTGTCAAAGCAGGTCGCAAGCTGCAAAACATATATGAGGCTTGTGTGATTGGCC 1500
Oy 501 ProGluAsnAspMetIleLeuGlyTyrLys 510
Db 1501 CCAAGCAATACATGATTTCTCCAGTACAAAG 1530

RESULT 14
ADS82001
ID ADS82001 standard; cDNA; 1533 BP.
XX
XX ADS82001;
XX
XX 18-NOV-2004 (first entry)
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
XX
XX Soybean, plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
XX Glycine max; line 29010CP01.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /product= "myo-inositol 1-phosphate synthase"
XX FT replace(260,G)
XX FT /tag= b
XX
XX US2003074685-A1.
XX

```

```

PD 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
PF
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
DR P-PSDB; ADS82002.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phylic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 11; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis for myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phylic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived
XX from seeds of a soybean plant (homozygous for one or more gene encoding a
XX mutant myo-inositol 1-phosphate synthase having decreased capacity for
XX the synthesis of myo-inositol 1-phosphate, where the gene confers a
XX heritable phenotype as mentioned above), and making or producing a
XX soybean protein product derived from seeds of a soybean plant with
XX heritable phenotype as mentioned above. The nucleic acid is useful for
XX altering raffinose saccharide, sucrose, phylic acid and inorganic
XX phosphate content of soybean seeds thus leading to valuable and useful
XX soybean products, since the presence of high concentration of raffinose
XX oligosaccharides in soy plants (and other legumes) can lead to flatulence
XX when consumed by humans. The present sequence encodes a mutant myo-
XX inositol 1-phosphate synthase.
XX
XX SO Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,596-249 Length: 1533
XX Score: 2593.00 Matches: 502
XX Percent Similarity: 99.22% Conservative: 4
XX Best Local Similarity: 98.43% Mismatches: 4
XX Query Match: 98.52% Indels: 0
XX DB: 13 Gaps: 0
XX
XX US-10-718-952-2 (1-510) x ADS82001 (1-1533)
XX
Oy 1 MetPheIleGlyAsnPheLysValGlyCysProAsnValLysTyrThrGluThrGluIle 20
Db 1 ATGTTCAATCGAGAAATTTTAAAGGTAGAGTCTTAATGTGAATACACGAGCTAGATT 60
Oy 21 GlnSerValTyrAsnTyrGlyThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 61 CAGTCGCGTACACTACGAAACCAACCACTTTTTCACGAGAACAGAAATGCACTAT 120
Oy 41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60

```

```

|||||
Db 121 CAGTGATGATGCAACCAATCCGTCACATACCAATTTAAACCAACCAACCATGTTCCA 180
Qy 61 LysLeuGlyValMetLeuValGlyTrpGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AAATTGGGGGTGATGCTTGTGGGTGGGGAACCAACCGCTCTACCCCTCACCGGTGGT 240
Qy 81 ValIleAlaAsnArgGlyIleSerTrpAlaThrLysAspLysIleGlnIleAsn 100
Db 241 GTTATTGCTACAGAGAGACATTTCAATGGGCTACAAAGAGCAAGATTCAACAGCAAT 300
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTTTGGGCTCCCTCACCAAGCCTCAGCTATTCAGTTGATCTTCCACGGAGAGAA 360
Qy 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ACTATAGCCCATTCACAGAGTCTGCTTCCAAATGTTAATCTGACGACATGTTGGG 420
Qy 141 GYTTPAAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GGAATGGGAATATCGACAACATGAACCTGGCTGATGCCATGGCCAGGCAAGGTTTGAC 480
Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
Db 481 ATCGATTTCAGAGACAGATGTTGAGGCTTACATGGAATCCATGGTTCCACTCCCGAATC 540
Qy 181 TyrAspProAspPheIleAlaIleAsnGlnGlyLysArgAlaAsnValIleLysGly 200
Db 541 TACGACCCGGAATTCATTTGCTGCCAACAGAGAGAGGTGCCAACACAGGATTAAGGGC 600
Qy 201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysValPheLysGlyAlaThr 220
Db 601 ACAAAACAGAGAGCAATGTCAGCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAGACACC 660
Qy 221 LysValAspLysValValIleLeuTrpThrAlaAsnThrGlyLysArgAsnLeuVal 240
Db 661 AAAGTGGACAGAGTGTGTTGCTCTGTGACCTGCCAACAGAGGATGATGCAATTTGGTT 720
Qy 241 ValGlyLeuAsnAspThrMetGlyAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
Db 721 GTAGGCTTAATGACACCATGAGAAATCTTGGCTGTGTGACAGAAATGAGCTGAG 780
Qy 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 AATTCCTCTCCACCTTGATGTCATGCTGTGTATGGAATAATGCTTTCATTAAT 840
Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAGCCCTCAGAACACTTTTGTACAGGGCTGATGATCTTCCATCGGAGAGAACACT 900
Qy 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTGGTGGAGATGACTTCAAGAGTGTGCAGACCAAAATGAATCTGTGTTGAT 960
Qy 321 PheLeuValGlyAlaGlyLysPheProThrSerIleValSerTyrAsnIleLeuGlyAsn 340
Db 961 TTTCTTGGGGGCTGGTGTCAAGCCAAATCATATGATTAGTTAACCAACATCTGGGAAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysSer 360
Db 1021 AATGATGATGATGATCTTCGGCTCCAAACCTTCGCTCCAAAGAAATCTCCAAAGAG 1080
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluIlePro 380
Db 1081 AACGTTGTTGACATATGTCACAGCAACCAATGCCATCTGTATAGGCTGTGTAACATCCC 1140
Qy 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCAATGTTGTTATTAAGATATGCTTACGTGGGGATGCAAGAGACCAATGAGAT 1200
Qy 401 GATYrThSerGluLePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
|||||

```

```

Db 1201 GAGTACACTTCAGAGATATTTCATGGGTGGAAAGAACACCATTTGTTTGACACACATGT 1260
Qy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGGATTCCTCTTACTCTCTCTATATCTTGGACCTTGTCTTCTTGTGAGCTGAGC 1320
Qy 441 ThrArgIleGluPheLysAlaGluAsnGlyGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAGAAATCAGTTTAAAGCTGAAATAGAGAAATTCACATTCATCCACCATTTGCT 1380
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATTCTAGGATATGACCAAGGCTCCTGTGTTCCACCGGATACACAGTGTGAAT 1440
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCATTGTCAAAGCAGCGTCAATGCTGGAACCATATATGAGGCTTGTGATTTGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
Db 1501 CCAGAGATTAACATGATCTCGAGTACAAAG 1530

RESULT 15
ID AAX90402
ID AAX90402 standard; cDNA to mRNA; 1950 BP.
XX AAX90402;
AC
XX 24-SEP-1999 (first entry)
DT
XX
XX Nicotiana paniculata INPS encoding cDNA.
DE
XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
KW water stress; resistance; de.
XX
XX Nicotiana paniculata.
OS
XX
XX JP1187879-A.
PN
XX
XX 13-JUL-1999.
PD
XX
XX 26-DEC-1997; 97JP-00359773.
PF
XX
XX 26-DEC-1997; 97JP-00359773.
PR
XX
XX (NINS) JAPAN TOBACCO INC.
PA
XX
XX WPI; 1999-451546/38.
DR P-PSDB; AAY24477.
XX
XX New INPS gene derived from Nicotiana genus plant - useful for conferring
PT resistance to water stress to plants.
XX
XX Claim 2; Page 6-8; 8pp; Japanese.
PS
XX
XX The present sequence encodes Nicotiana paniculata inositol monophosphate
CC synthase (INPS), designated NpINPS1. INPS can be used to confer water
CC stress resistance to a plant
XX
SQ Sequence 1950 BP; 570 A; 377 C; 425 G; 578 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,93e-237 Length: 1950
Score: 2472.00 Matches: 470
Percent Similarity: 96.864 Conservative: 24
Best Local Similarity: 92.164 Mismatches: 16
Query Match: 93.928 Indels: 0
DB: Gaps: 0

US-10-718-952-2 (1-510) x AAX90402 (1-1950)
Qy 1 MetPheIleGluAsnPheLysValGlyCysProAsnValLysTyrThrGluThrGluIle 20
|||||

```

```

Db      92 ATGTTATTGAGAAATTTTAAGGTGAGACCCCAACGTTAAGTACACCGAAAGTGAAT 151
QY      21 GlnSerValTyrAsnTyrGlnThrGluLeuValHieGluAsnArgAsnGlyThrTyr 40
Db      152 CACTGTGTTATGATTTATCAAAACCACTGATGTTAGTTTCATGATGACAGAAAATGGACAT 211
QY      41 GlnTrrPileValIleAspProLysSerValIleTyrGluPheLysThrAsnIleHisValPro 60
Db      212 CAATGGACCGTCAAGCCTAAGACTGTCAAATATGAGTTCAAGACTGATTCATGTTCC 271
QY      61 LysLeuGlyValMetLeuValGlyTrrPglYglYAsnAsnGlySerThrLeuThrGlyGly 80
Db      212 AATTATGGGGGTATGTTGTTGATGGGATGGAAACAAATGTTCAACCTTGACCGGTGT 331
QY      81 ValIleIleAsnArgGluGlyIleSerTrrPalaThrLysAspLysIleGlnGlnIleAsn 100
Db      332 GTTATTGCTTACACGAGAAAGAAATTCATGGGCCCAAAAGATTAAGGTGCAACAGCCAT 391
QY      101 TrrPheGlySerLeuThrGlnIleSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db      392 TACTTTGGCTCTTACTACTCAGGCTTACTATTCGAGTTGGGCTTTCAATGGAGAGAG 451
QY      121 IleTrrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db      452 ATCTATGCTCCATTTAAAGCCCTCTTCCAAATGGTCAATCCAGATGACGTAGTGTGGA 511
QY      141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db      512 GGATGGGACATACGCAACATGAATTTAGCAGATGCGATGGCCAGGCTTAAGGATTTGAT 571
QY      161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProGluIle 180
Db      572 ATTGATCTACAAAGAGAGTTGAGGCCCTCAATGATCTATGATCCACTCCCTGGTATC 631
QY      181 TrrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
Db      632 TATGACCTGATTTCTCTGCTGTAACCAAGGCTCAGTCCAAACACGTGATCAAAAGGA 691
QY      201 ThrLysGlnGlnValGlnIleIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db      692 ACCAAGAAAGAAACAAATTCATCAATCATTAAGATTTAGGAGTTTAAGAAAGAAAC 751
QY      221 LysValAspLysValValIleuTrrPrrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      752 AAGTGGACAAAGTGTGATGATGATGCTGCACTGAACATGAAGTAAATGTGTT 811
QY      241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluIleGlu 260
Db      812 GTTGAATTATGACATATGAAACCTCTTGTCTGTGACAGAAATGAAAGCTGAA 871
QY      261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      872 ATATCTCCTTCCACTTGTATGCTATGCTGATTCCTGAAAATGTGCTTTATTAAT 931
QY      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      932 GGAAGCCCCAGAAACACCTTTGTCAGGACCTCATGATTTGGCCATCAAGAAACACA 991
QY      301 LeuIleGlyIleAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      992 TTGATTGGTGGTATGACTTTAAGATGTGTCAAAACCAAAATGAAGTCAGTGTGTTGAT 1051
QY      321 PheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db      1052 TTCCTTGTGGAGCTGTATTAAGCCAACTCAATTTGTAGCTACAAACATTTGGGTAAC 1111
QY      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db      1112 AATGATGAGATGAATCTGTCTGCCCTCAAACTTTCCGGTCAAAAGAGATCTGAAAAGT 1171
QY      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db      1172 AATGTTGTGATGACATGGTTTCAAAGCAATGCGATCTTATGAGCTTGAGAGACCCCT 1231

```

```

QY      381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1232 GACCATGTTGTTGATTAAGATATGTCATATGTGGAGACAGCAAGAGGCAATGAT 1291
QY      401 GlnTrrThrSerGlnIlePheMetGlyIleLysSerTrrIleValLeuHisAsnThrCys 420
Db      1292 GAGTACACATCTGAGATTTTCAATGGGGGAAAGAACACCATTTGTTTCCAAATACCTGT 1351
QY      421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1352 GAGGATTCATCTTTAGCTGCTCCATTAATTAATTGATTTGGTCTTCTGCTGAATCTCAGT 1411
QY      441 ThrArgIleGluPheLysAlaGluAsnGlyIleLysPheHisSerPheHisProValAla 460
Db      1412 ACCCGCATTCAGCTCAAGCTGAAGAGAGGAGGATGTTCCACTCTCCACCCCGTGGCT 1471
QY      461 ThrIleLeuSerTyrLeuThrThrAlaProIleIleLeuAspLeuValProProGlyThrProValValAsn 480
Db      1472 ACTATCTCAGCTACCTTACCAAGGCTCTCTGTGTACACCAAGTACACCAAGTGTGAT 1531
QY      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1532 GCACTCTCAAGCAGAGGGCAATGCTTGAGAAATATTTGAGGGCTTGTGTTGACTTGA 1591
QY      501 ProGluAsnAsnMetIleLeuGluTrrLys 510
Db      1592 CCAGAGAAACATCATGATTCCTGGAATPACAA 1621

```

Search completed: June 7, 2005, 17:47:34
 Job time : 677.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:22:56 ; Search time 219 Seconds
(without alignments)
3810.507 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MLEIKPKVSCPNVKTETETI.....NIMRACVGLAPENMILEYK 510

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODBI=frame+.p2n.model -DEV=xlp
-O=/cg2_1/USPTO.spool.p/US10718952/rnat.06062005.173401.12868/app.query.fasta_1.1294
-DB=Issued Patents NA -QFMT=fastp -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-NODE=LOCAL -OUTFMT=pct -NOMEXE -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952.OCN_1.141.0.rnat.06062005.173401.12868 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGIOLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cg2_6/prodata/1/ina/5A_COMB.seq:*
2: /cg2_6/prodata/1/ina/5B_COMB.seq:*
3: /cg2_6/prodata/1/ina/5A_COMB.seq:*
4: /cg2_6/prodata/1/ina/5B_COMB.seq:*
5: /cg2_6/prodata/1/ina/PCUS_COMB.seq:*
6: /cg2_6/prodata/1/ina/backfillseq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2336	88.8	1931	3	US-09-118-442-10 Sequence 10, Appl
2	2336	88.8	1931	3	US-09-677-064-10 Sequence 10, Appl
3	2336	88.7	1931	4	US-09-727-628-10 Sequence 1, Appl
4	1813	68.9	3546	3	US-09-118-442-15 Sequence 15, Appl
5	1813	68.9	3546	3	US-09-677-064-15 Sequence 15, Appl
6	1812	68.8	3546	3	US-09-118-442-14 Sequence 14, Appl
7	1812	68.8	3546	3	US-09-677-064-14 Sequence 14, Appl
8	1382.5	52.5	1602	4	US-09-734-237B-72 Sequence 72, Appl
9	1382.5	52.5	1605	4	US-09-734-237B-74 Sequence 74, Appl
10	1376.5	52.3	1578	4	US-09-248-796A-3131 Sequence 111, Ap
11	1220.5	46.4	77626	4	US-09-949-016-12608 Sequence 12608, A
12	1098.5	41.7	1231	3	US-09-397-787-34 Sequence 34, Appl

c	13	545	20.7	34316	4	US-09-902-540-1257	Sequence 1257, Ap
	14	541	20.6	1335	4	US-09-902-540-5256	Sequence 5256, Ap
	15	155	5.9	294	4	US-09-313-294A-4684	Sequence 4684, Ap
	16	142	5.4	42325	4	US-08-311-731A-131	Sequence 131, Ap
c	17	129.5	4.9	4403765	3	US-09-103-840A-2	Sequence 2, Appl
	18	128.5	4.9	4411529	3	US-09-103-840A-1	Sequence 1, Appl
c	19	115.5	4.4	7571	3	US-08-961-527-2	Sequence 2, Appl
	20	113.5	4.3	274	3	US-09-118-442-21	Sequence 21, Appl
	21	113.5	4.3	274	3	US-09-677-064-21	Sequence 21, Appl
	22	113.5	4.3	2190	4	US-09-107-433-2104	Sequence 2104, Ap
	23	113.5	4.3	2196	4	US-09-583-110-2059	Sequence 2059, Ap
	24	112.5	4.3	2529	4	US-09-489-039A-4399	Sequence 4399, Ap
	25	111.5	4.2	1359	4	US-09-248-796A-6542	Sequence 6542, Ap
	26	111.5	4.2	2193	1	US-09-014-897-1	Sequence 1, Appl
	27	111.5	4.2	2193	1	US-09-014-897-3	Sequence 3, Appl
	28	111.5	4.2	2193	1	US-08-731-716-1	Sequence 1, Appl
	29	111.5	4.2	2193	1	US-08-731-716-3	Sequence 3, Appl
	30	111	4.2	8287	4	US-09-023-655-1441	Sequence 1441, Ap
	31	109.5	4.2	2244	4	US-09-583-110-405	Sequence 405, Ap
	32	109.5	4.2	2256	4	US-09-107-433-772	Sequence 772, Ap
	33	109	4.1	3421	4	US-08-956-171E-337	Sequence 337, Ap
	34	109	4.1	3421	4	US-08-781-986A-337	Sequence 337, Ap
	35	109	4.1	8040	1	US-08-596-291-1	Sequence 1, Appl
	36	109	4.1	8040	1	US-09-100-804-1	Sequence 1, Appl
	37	109	4.1	8043	5	PCT-US94-09943-1	Sequence 1, Appl
	38	109	4.1	8119	3	US-09-290-640-45	Sequence 45, Appl
	39	109	4.1	8119	3	US-09-665-615B-45	Sequence 45, Appl
	40	108	4.1	1052	3	US-08-961-527-244	Sequence 244, Ap
	41	107.5	4.1	5892	4	US-09-583-110-2582	Sequence 2582, Ap
	42	107.5	4.1	5904	4	US-09-107-433-2280	Sequence 2280, Ap
	43	106.5	4.0	2337	4	US-09-489-039A-1284	Sequence 1284, Ap
	44	106	4.0	1620	4	US-09-543-681A-2705	Sequence 2705, Ap
	45	106	4.0	3729	4	US-09-107-532A-1587	Sequence 1587, Ap

ALIGNMENTS

RESULT 1
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10
Alignment Scores:
Pred. No.: 1.05e-283
Score: 2336.00
Percent Similarity: 93.33%
Best Local Similarity: 87.25%
Query Match: 88.75%
DB: 3
Length: 1931
Matches: 445
Conservative: 31
Mismatch: 34
Indels: 0
Gaps: 0

US-10-718-952-2 (1-510) x US-09-118-442-10 (1-1931)

```

QY 1 MetPheIleIeGluAsnPhenylsValIGlucySProAsnValLysTyThrGluThrGluIle 20
Db 99 AAGTTCATCGAGAGCTTCGCGTCAAGACCCCGTGGTACGGCCGATGAGATC 158
QY 21 GlnSerValTYrAsnTYrGluThrThrgluValHisGluAsnArgAsnGlyThrTYr 40
Db 159 GAGTCGAGTACCGGTACGACGACGAGCTGTHACAGAGGCAAGAGCGGCCCTCA 218
QY 41 GlnTPriLeValIleValProLysSerValIlyTYrGluPheLysThraSnlIleHisValPro 60
Db 219 CGCTGGGTGTCGCCGCCAAGTCCGTCAGTCAACTTCGAGACCAAGCCCGTCCCC 278
QY 61 LysLeuGluValMetLeuValIGlYTPriGlyValAsnAsnGlySerThrLeuThrgly 80
Db 279 AAGCTCGGGGTATGCTGTGGGTGGGAGGCAACAGGGGTCCACGCTGACGGCTGGG 338
QY 81 ValIleAlaAsnArgGluGluIlySerTPriAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 339 GTTCATTCGCCACAGAGGAGGAGTCTCATGGCGACCAAGGACAGAGTGCAGACCAAC 398
QY 101 TYrPheGlySerLeuThrglnAlaSerAlaIleArgValIGlySerPheGlnIGlyGlu 120
Db 399 TACTACGGCTCCCTCACCACGCTCCACATCAGAGTCCGACGCTCAACCGGGAGAG 458
QY 121 IleTYrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 459 ATCTATGCGCCCTTCAGAGGCTCTCTCCATAGTGAACCCAGACACACTTGTGTGGGA 518
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 519 GGCTGGGACATTAAGCAACATGAACCTGGCCGACTCATACCAAGGCCCAAGGTGCTGAT 578
QY 161 IleAspLeuGlnLysGlnLeuArgProTYrMetGluSerMetLeuProLeuProGlyIle 180
Db 579 ATTGACCTGCAGAGAGAGCTCAGGCCCTCATGAGTGCATGATGATGATGATGATGAT 638
QY 181 TYrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
Db 639 TATGATCCGAGCTTCATCGCGCTAACCGAGCTCTCGGCCCAACAGTCTCATCAAGGGC 698
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db 699 ACCAAGAAGAACAAGGTGAGCAGATCATCAAGATACAGGAGTTTAAGAGAAAGAAC 758
QY 221 LysValAspLysValValIleuTYrPThrAlaAsnThrgluArgTYrSerAsnLeuVal 240
Db 759 AAAGTCGACAAAGATAGTGTGTGTGACTGCACAACTGAAGAGTATAGCAATGTGTGC 818
QY 241 ValIGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGlnAlaGlu 260
Db 819 GCTGGTCTAACGACAGATGAGATCTACTGGCTGTGTGACAGAAAGAGAGGGAG 878
QY 261 IleSerProSerThraTYrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 879 GTATCCCATCAACATATATGCAATGCTGTGTATATGGAAGGGGTGCGCTTCATCAT 938
QY 281 GlySerProGlnAsnThraPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 939 GGGAGGCCCCCAAGACCTTTGTGCTGGGCTGATGTGATCTGTCTATAAAAAACAACCTGC 998
QY 301 LeuIleGlyValAspAspPheLysSerGlyGlnThryMetLysSerValLeuValAsp 320
Db 999 TTGATTTGGTGTGACACTTCAAGAGTGAACAGACCAAAATGAATCTTCTTGTGCTGAT 1058
QY 321 PheLeuValGlyValAGlyIleLysProThrSerIleValSerTYrAsnHisLeuGlyAsn 340
Db 1059 TTCCTTTGTTGGTGTGATTAAGCCCACTCATGTATAGCTACCAACCACTTGGGAAAC 1118
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360

```

```

Db 1119 AACGATGCGATGAACCTGTCTGCCCTTCMAACATTCAGGTCCAAAGAGATCTCCAAAGAC 1178
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTYrGluProGlyGlnHisPro 380
Db 1179 AAGCTGTGATGATGATGATGATCTTCAGACAAATGCCATCTCTATGAGCCCGGAGAGATCCC 1238
QY 381 AspHisValValIleLysTYrValProTYrValGlyAspSerLysArgAlaMetAsp 400
Db 1239 GATCATGTGTTGTTCATCAAGATATGTCGTAACGGGAGAGACAGACAGAGGGCTATGAGAC 1298
QY 401 GlnTYrThraSerGluIlePheMetGlyGlyLysSerThriLeuValLeuHisAsnThryS 420
Db 1299 GAGTACACCTCAGAGATTTTATGAGCGGCAAGAACACCATTCGTGTGCACAAACCTGT 1358
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1359 GAGGACTGCTCTCGCCGACCTTATCTTGTATCTGTGTCTTGTGGCTGAGCTGAGC 1418
QY 441 ThraGlnGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db 1419 ACCAGATTCACAGCTGAAGCTGAGGAGAGAGACAAATTCACATCTTCCACCCGGTGGCC 1478
QY 461 ThrIleLeuSerTYrLeuThryAlaProLeuValProProGlyTYrProValValAsn 480
Db 1479 ACCATCTTGAGTTACTTCAACAGGACCCCTGTCTCCCTGGACACCGGTGGTGAAC 1538
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValIGlyLeuAla 500
Db 1539 GCTCTGGCCAAACAGAGGCGCATGCTGGAGAAACATCATGAGGGCTGCGTGGCTGGCC 1598
QY 501 ProGluAsnAsnMetIleLeuGluTYrLys 510
Db 1599 CCAGAGAACACATGATCTTGAGATCAAG 1628

```

RESULT 2

```

US-09-677-064-10
/ Sequence 10, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Calt, Susan J.
/ APPLICANT: Beach, Larry R.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677,064
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 10
/ LENGTH: 1931
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-677-064-10

```

```

Alignment Scores:
Pred. No.: 1,056-283 Length: 1931
Score: 2336.00 Matches: 445
Percent Similarity: 93.33% Conservative: 31
Beet Local Similarity: 87.25% Mismatches: 34
Query Match: 88.75% Indels: 0
DB: 3 Gaps: 0

```

US-10-718-952-2 (1-510) x US-09-677-064-10 (1-1931)

```
QY 1 MetPhe11eGluAanPhe1ySValGluCySProAanVal1yTyrThrGluThrGlu1e 20
Db 99 ATGTTTCATGAGAGAGCTTCGGCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 158
QY 21 GlnSerVal1TyAanTyGluThrThrGluLeuVal1H1eGluAanArgAnGlyThrTy 40
Db 159 GAGTCGAGATACCGGATACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 218
QY 41 GlnTP11eVal1ySPro1ySVal1yTyrGluPhe1ySVal1yTyrGluPhe1ySVal1y 60
Db 219 CGCTGGTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 278
QY 61 LysLeuG1yValMetLeuVal1yTyrP1yG1yVal1ySAnG1ySVal1yTyrGlu1y 80
Db 279 AAGCTCGGGGATGCTTGTGGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338
QY 81 Val11eVal1yAanArgGluG1y11eSerTP1a1aThrLysAsp1yS11eGln1a1aAn 100
Db 339 GTCATTCGCAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398
QY 101 TyrPheG1ySerLeuThrGln1a1aSer1a1a1eArgVal1y1ySerPheGln1yGlu 120
Db 399 TACTACGGCTCCCTCCACCGCTCCACCATCAGAGTGGCGAGCTTACAGAGAGAG 458
QY 121 11eTyR1a1aProPhe1ySVal1yProMetVal1aAnProAsp11eVal1yPheG1y 140
Db 459 ATCTATGCCCGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
QY 141 G1yTP1a1a1eSerAanMetAanLeu1a1aAsp1a1aMet1a1aArg1a1aVal1yPhe 160
Db 519 GGCTGGGACATTAAGCAATGAACCTGGCGAGCTCATGACCAAGGAGAGAGAGAG 578
QY 161 11eAspLeuGln1ySVal1yLeuArgPro1yTyrMetGluSerMetLeuPro1yG1y1e 180
Db 579 ATTTGACCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638
QY 181 TyrAspProAspPhe11eVal1a1aAnGln1yGlu1yAArg1a1aAnAanVal11eTySg1y 200
Db 639 TATGATCCGAGATTCATCCGCGCTAACAGAGGCTCTCGGCGCAACAGTGTATCAAGAGG 698
QY 201 ThrTySglnGln1yA1a1aGln1y11e1ySAsp11eLysA1a1aPhe1ySgln1a1aThr 220
Db 699 ACCAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 758
QY 221 LysVal1a1aPhe1yVal1yVal1yLeuTP1a1a1aThrGluArgTySerAanLeuVal 240
Db 759 AAAGTGAGACAGATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 818
QY 241 ValG1yLeuAanAspThrMetGluAanLeuVal1a1aVal1a1aArgAnG1yGlu1y 260
Db 819 GCTGGTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
QY 261 11eSerProSerThrLeuTyR1a1a1eVal1a1aCySVal1yMetGluAanVal1yProPhe11eAan 280
Db 879 GATATCAACATCAACATATATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 938
QY 281 G1ySerProGlnAanThrPheVal1ProG1yLeu11eAspLeuVal11eVal1a1aArgAnThr 300
Db 939 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998
QY 301 Leu11eG1y1yAspAspPhe1ySVal1yGlnThyMetLeuSVal1yLeuVal1a1aP 320
Db 999 TTGATTTGGTGTGACATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1058
QY 321 PheLeuVal1y1a1aG1y11eLysProThrSer11eVal1ySerTyRAn11eSgln1yAan 340
Db 1059 TTCCTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1118
QY 341 AaAnAspG1yMetAanLeuSer1a1aProGlnThrPheArgSerTySgln1ySerTySg 360
Db 1119 AACGATGGGATGAACCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1178
QY 361 AaAnVal1a1aAspAspMetVal1aAnSerAanVal11eLeuTyRglnProG1yGln1ySPro 380
```

```
Db 1179 AAGTGTGATGACATGATCTCGAGCATGCAATCTCTATAGAGCCGAGAGATGCC 1238
QY 381 AspH1sVal1yVal11eLysTyRVal1ProTyRVal1yAspSerTySVal1a1aMetAap 400
Db 1239 GATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1298
QY 401 GlnTyRThrSerGln11ePheMetG1y1ySVal1ySVal1yVal1yLeuTh1a1aAnThrCyS 420
Db 1299 GAGTACACTTCAGAGATCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1358
QY 421 GluAspSerLeuLeuVal1a1aPro11e11eLeuAspLeuVal1yLeuLeuVal1ySer 440
Db 1359 GAGAGATCGCTCTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1418
QY 441 ThrArg11eGluPhe1yVal1a1aGluAanG1y1ySVal1ySVal1ySVal1ySVal1y 460
Db 1419 ACCAGAGATTCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1478
QY 461 Thr11eLeuSerTyRLeuThrTyR1ySVal1a1aProLeuVal1ProProG1yThrProVal1yAan 480
Db 1479 ACCATCTTGAATTAATTCACCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1538
QY 481 AlaLeuSerTySgln1yAArg1a1aMetLeuGln1yAan11eMetArg1a1aCySVal1yLeuVal 500
Db 1539 GCTTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1598
QY 501 ProGluAanMet11eLeuGlnTyR1yS 510
Db 1599 CCAGAGAACATGATCTTGAATTAACAG 1628

RESULT 3
US-09-727-628-1
; Sequence 1, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727,628
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,612
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)..(1699)
US-09-727-628-1

Alignment Scores:
Pred. No.: 1,44e-263 Length: 1959
Score: 2335.00 Matches: 447
Percent Similarity: 93.33% Conservative: 29
Best Local Similarity: 87.65% Mismatches: 34
Query Match: 88.72% Indels: 0
Gaps: 0

US-10-718-952-2 (1-510) x US-09-727-628-1 (1-1959)

QY 1 MetPhe11eGluAanPhe1ySValGluCySProAanVal1yTyrThrGluThrGlu1e 20
Db 137 ATGTTTCATGAGAGAGCTTCGGCGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196
QY 21 GlnSerVal1TyAanTyGluThrThrGluLeuVal1H1eGluAanArgAnGlyThrTy 40
```

```

Db      197 GAGTCGAGTACCGGTACACAGACGAGCTGTGTCACGAGGCCAAGAGCGCGCTCC 256
      41 GINTPILlEaLVaSPheProLysSerValLySTyrgLuphLeThraSnLlEhISaLVaPro 60
      257 CGCTGGGTGTCGCGCCCAAGTCCGTCAAGTCAACTTCGAGCAAGCAAGCGCGGCTCCC 316
      61 LysLeuGLyValMetLeuValGLyTPGlyGLYAsnaSngLYSerThrLeuthrGLyGLY 80
      317 AAGCTCGGGGTATGCTTGTGGGGTGGGAGGCAACACGGGTCCACGCTGACGGTGGG 376
      81 ValLLlEaLVaSPheProLysSerValLySTyrgLuphLeThraSnLlEhISaLVaPro 100
      377 GTCATGTCACAAAGGAGGAGATCTCATGGGCGAACCAAGAGCAAGGTGACAGCAAGCAAC 436
      101 TyrrhEgLYSerLeuThrLnaLaseRaLLlEaLVaGLYSerThrheGLnGLYGLn 120
      437 TACTACGGGTCCCTCCAGGCTTCACATCAGAGTGGCAGCTCAACAAAGGGAGGAG 496
      121 lIeTYrLaProPhelYSerLeuLeuProMetValaSPheProLysSerValLySTyrgL 140
      497 ATATATGCCCGCTTCAAGAGCTCTCAACCAAGTGAACCAAGCAAGCTTGTGTGGA 556
      141 GLYTPASPILeserAmeThraSnaLeuValaSPheProLysSerValLySTyrgL 160
      557 GCGTGGACATCAGCAGCATGAACCTGGCAGATGCGATGACCAAGGCGCAAGGCTGAC 616
      161 lIeAPheLGLnGLnLeuValProTYrMetGLYSerMetLeuProLysSerValLySTy 180
      617 ATTGACCTCAGAAAGAGCTCAGGCTCTACATGATGATGATGATGATGATGATGATG 676
      181 TYrASPProASPheLlEaLVaSPheProLysSerValLySTyrgL 200
      677 TATGATCCCGATCTTATCCCGCTTACCAAGGCTCTCGTCCCAACATGTCATCAAGGCG 736
      201 ThLYeGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLn 220
      737 ACCAAGAAAGAAAGAGTGGAGCATCATCAAAAGATACAGGAGGATTAAGGAGAAAG 796
      221 LysValaSPheLysValaLVaLLeuTPThraLaSPheThGLnGLnGLnGLnGLnGLn 240
      797 AAAGTGAAGAGGTATGCTGTGCTGAGCTGCAAAACATGAAAGTACAGCAATGTATGT 856
      241 ValGLYLeuAsnaSPheThraSnaLeuLeuValaLVaLAspaRgaSngLlEaLVa 260
      857 GCTGCTCTCACACACATGAGATCTGCTGTGATCTGTGCAAGAACAGAGGCGGAG 916
      261 lIeSerProSerThrLeuTYrLaLLlEaLVaCYsValMetGLnSnaValProPhelLea 280
      917 ATCTCGCATCAACATATATGCTGCTGTGTCACGAGGAGGCTGCGCTCATCAAT 976
      281 GLYSerProGLnSnaThraSnaLeuValProGLYLeuLlEaLVaSPheThraSna 300
      977 GGGAGGCCCCAGAAACCTTTGTGCTGGCTGATGATCTTGTCAAAAGCAACCTGC 1036
      301 LeuLLlEaLVaSPheProLysSerValLySTyrgLuphLeThraSnLlEhISaLVa 320
      1037 CTGATCGGTGTGTCACATCTTCAAGAGTGGGAGCAACCAAGTGAATCGGTCTGTGAT 1096
      321 PhELeuValGLYValGLYLeuSProThraSnaLlEaLVaLYSerThraSnaSleuGLY 340
      1097 TTTCTTGTGTGCTGGAATTAAGCCACCTCGATGTGAGCTAACCAACCTGGGAAAC 1156
      341 AaNaSPGLYMetAmeThraSnaLeuValProGLnThraSnaSleuGLYLeuSPro 360
      1157 AACGAGGAGTAACTGTCTGCTCCCTCAAACTTACAGTCCAAAGAGATCTCAAGAGC 1216
      361 AaNaValaLVaSPheProLysSerValLySTyrgLuphLeThraSnLlEhISaLVa 380
      1217 AACGTGTGATACATGCTCTCAAGCAATGCAATCTCTATGAGGCGCGGAGCATCCC 1276
      381 AaPhISaLVaLVaLlEaLVaProTYrValProTYrValGLYAspSerLYaRgaLMeTa 400

```

```

Db      1277 GATCATGTTGTCATCAAGTATGTCGCTGATGTGGAGACAGTAAAGGCTATGAC 1336
      401 GlnTYrThraSnaLlEaLVaSPheProLysSerValLySTyrgLuphLeThraSn 420
      1337 GAGTACATCAAGATCTTCAAGGCGGCAAGACCATGCTGTGTCACAAACCTGCC 1396
      421 GlnaSPheLeuValaLVaProLlEaLVaSPheLeuValaLVaLLeuValaLVaL 440
      1397 GAGACCTGCTCTCGCGGACCGATCATCTCGATCTGTGCTCTGCTGTGAGCTCAGC 1456
      441 ThraGLlEaLVaSPheProLysSerValLySTyrgLuphLeThraSnLlEaLVa 460
      1457 ACCAGATTCAGTAAACCTGAGGAAACGACAAAGTTCATCTTCAACCGGAGGCTC 1516
      461 ThrlEaSerTYrLeuThraSnaLlEaLVaProLysSerValLySTyrgLuphLe 480
      1517 ACCATCTTACCTACCTCAAGGACCATCTGTTCAACCGGAGCTGTGAGTGAAC 1576
      481 AlaLeuSerLYeGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLnGLn 500
      1577 GCTCTTGCAAGCAGAGGCGATGCTGGAGAACATCATGAGGCTTGGCTGGCTGCC 1636
      501 ProGLnSnaMetLlEaLVaSPheProLysSerValLySTyrgLuphLeThraSn 510
      1637 CCAGAAACAAATGATCTCTGAGTAAAG 1666

RESULT 4
US-09-118-442-15
/ Sequence 15, Application US/09118442B
/ Patent No. 6197561
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ APPLICANT: Wang, Xun
/ APPLICANT: Bowen, Benjamin A.
/ TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
/ FILE OF INVENTION: Plants and Uses Thereof
/ FILE REFERENCE: 0706
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ EARLIER FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-118-442-15

Alignment Scores:
Pred. No.: 2,71e-217 Length: 3546
Score: 1813.00 Matches: 443
Percent Similarity: 40.86% Conservative: 31
Best Local Similarity: 38.15% Mismatches: 36
Query Match: 68.88% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-2 (1-510) x US-09-118-442-15 (1-3546)
      1 MetPhelLlEaLVaSPheProLysSerValLySTyrgLuphLeThraSnLlEhISaLVa 20
      60 ATGTCATCAAGAGTTCGCGCTGAGAGGCGCCAGTGGCTGAGCGCGGAGATCA 119
      21 GlnSerValTYrAsnTYrGLnThraSnaLlEaLVaLlEaLVaSPheProLysSerValLySTy 40
      120 GAGTCGAGTACCGGTACACAGACGAGCTGTGTACAGAGGCAAGAGCGCGCTCA 179

```

QY 41 GlnTrpIleValIysPheIysSerValIysIleTyrgluPheIysPheIysPheIysIleHisValPro 60
Db 180 CGCTGGTGTGTCGCCCAAGTCCGTCAAGTACAACTTCGAGACCAAGCCCGCTCC 239
QY 61 IysIleuGly----- 63
Db 240 AAGCTCGGGTATGTACGAGATGACGGCCCTACCTCAGCTCTGTGAAACCTCTCTC 299
QY 63 ----- 63
Db 300 CGGTGCTCAGTCAAAATCTCCGTGAGATCACTGGTCGCGCTTCCCTCTAAATCTAA 359
QY 63 ----- 63
Db 360 TGAATAATTCTACTGCTTGGCTGAGACGAAACCGTGTATTTGTGACAGCTACGACAC 419
QY 63 ----- 63
Db 420 ACTTGCCCATCGGATCGGTCAAAATGAGCTCGATTGAAATTGATTCGATGGTCCCTT 479
QY 63 ----- 63
Db 480 TTCCATATTTGATCATCTCCCTGCTACTGTGCAATGATTAACAGAAAGTCCCTTCTC 539
QY 63 ----- 63
Db 540 TGAATTGTTGTAGGCTTTTGTCTGTGACGTGAGCTGTATCAATTTGTTCATGTA 599
QY 63 ----- 63
Db 600 AGATCAAAATTCACGACGAGACATGACGACAGACAGAACTCATTAACCTAGCAATTGA 659
QY 63 ----- 63
Db 660 TACTAGATTACTGCAAGTGTGATACGCGCAATCTGCATCTGACCCCTTTGTT 719
QY 63 ----- 63
Db 720 TAATTCCTGTCTTATGATGTTGCTTACGTGACGCTGTGTGTATGTTGTCAGGC 779
QY 63 ----- 63
Db 780 TGTGACCGCTGTGTCTGTGTCGACGATGATGCCAACTTTCTGTTCTGTGTGTCAGG 839
QY 64 ValMetIleuValIyTrpGlyIysAsnGlySerThrIleuThrGlyValIleAla 83
Db 840 GTGATGCTTGTGTGGGTGGGAGGCAACAGGGTCCACGCTGACGCGTGGGTCAATTGCC 899
QY 84 AsnArgGlu----- 86
Db 900 ACCAGGAGTGAAGTACTTAATTTGCTATATTGCTTTCGTTGTTTCAGTTATTA 959
QY 87 ----- 92
Db 960 ATGCGCTGACAGAACTGAATTTTGTGTGGCTGTTCAAGGGATCTCATGCGCAGC 1019
QY 93 IysAspIysIleGlnIleAsnTyrrPheGlySerIleuThrGlnIleSerAlaIleArg 112
Db 1020 AAGGACAAAGTGCAGACCAACTACTACGGCTC-CTCACCCAGGCTCCACCATCAGA 1078
QY 113 ValIysSerPheGlnIyGluGluIleTyrrAlaProPheIysSerIleuProMet--- 131
Db 1079 GTCGGAGTCAACAGGGAGAGATCTATGCGCTTCAGAGCTCTCTCCCATGGTA 1138
QY 131 ----- 131
Db 1139 ATCTATATAGACTTGAATAACTCTCTTTTACTGAACCAACATACATAAAG 1198
QY 131 ----- 131
Db 1199 CATATTCGTAAGTGTCTAGTTGATGTATAAAGTAACCTGCTTTCAGGCGAGTGTTC 1258
QY 131 ----- 131

Db 1259 TCAAGTAACGGAATGTTAATCATTTGGGTGAAAAACAAAGTTCTAATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAATGTTAACTTAGCATATGAAAAAGGAGACATTGTGAAGAAAGTGTGAAACATTC 1378
QY 131 ----- 131
Db 1379 GACTGGTCTGCATGTGTGTATCTTACCTGCAAGTCAAAAGTTCGTGTGTTAGCTCA 1438
QY 131 ----- 131
Db 1439 AGGTTCAGCATTTTGAATTACACTCGTGCAGTATTGACATGTGTCTAAGTGTGCA 1498
QY 131 ----- 131
Db 1499 GATTGCAAGCTCGGTTGTTATCTTTCATGACCAAGTGTGAACGTGTTTCA 1558
QY 132 -ValAsnProAspAspIleValPheGlyGlyTrpAspIleSerAsnMetAsnLeuAla 151
Db 1559 GGTGAACCCAGACGATTTGTGTGAGGCTGGACATTGACCAATGAACTGACCTGCCGA 1618
QY 151 PalMetAlaArgAlaIysValPheAspIleAspIleuGlnIysGlnIleuArgProTyrme 171
Db 1619 CTCATGACCAAGGCGCAAGTCTGATATTAACCTGCAAGACAGCTCAGGCTTCAAT 1678
QY 171 GluSerMetIleuProIleu-ProGlyIleTyrrAspProAspPheIleAlaAsnGlnG 191
Db 1679 GGAATCATGATGTCACCTTCCCGGATCTAATGATCCGACCTTCAATCCGGCTAACAG 1738
QY 191 IugIuArgAlaAsnAsnValIleIyGlyThrIyGlnIugIuValGlnGlnIleIle 211
Db 1739 GCTCTCGCCCAACAGTGCATCAAGGCGCACCAAGAAAGACAGGAGAGATCATCA 1798
QY 211 ValAspIle----- 213
Db 1799 AGGATATCAGTATATGATATGATGCTAACGTGCTGTGCTAAGGTGACCCAGTG 1858
QY 213 ----- 213
Db 1859 CAACCTAAACAAATTAATACTACTATGAAATTTGGTAATATACATATCAGAACAT 1918
QY 213 ----- 213
Db 1919 ATGTTTAAACCGTGCACTTAGAAGTCTGCATGTATGTTGGACAATTTGACATTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGCTCACTTGATGAGACTCCACAAAGAACTAAACTACTGAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTCCGATATGATGATTTTGTGTGACATGTTGAAGATCTAGATTAAAGT 2098
QY 213 ----- 213
Db 2099 GGTGAAGAAATATGTTCACTAGCATAGTAATCATTAACAGACTATGCTTATGTA 2158
QY 214 -----LysAlaPheIyGlnIleThrIyValAspIysValIleValIleuTrpT 230
Db 2159 GCTAATGAACAGGAGTTTAAAGAGAGAACAAAGTGAAGAATGATGTTGTGTGGA 2218
QY 230 hrAlAsnThrGluArgTyrrSerAsnLeuValIleGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACCTGAAGAGTATAGCAATGTGTGCGTGTCTCAACGACAGATGGAATTC 2278
QY 250 euLeuAlaIleValIlePheArgAsnIleValIleGlyIleSerProSerThrIleuTyrrAlle 270
Db 2279 TACTGCACTGTGTGACAAAGACGAGCGAGGTATCAACATCAACATATATGTCATTG 2338
QY 270 IacYeuValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro- 289

```

Db 2339 CCTGTGTCATGAGAGGGGTCGCTTCATCATGAATGGAGAGCCCGAAGACACCTTTGTGCTG 2398
Qy 289 ----- 289
Db 2399 GTGCGGTGTTGGTGTGTTGCAAAAAGCTTCATGCTGTGCATTTCTGTTCCAAAAGTTTC 2458
Qy 290 -----GlyLeuIle 292
Db 2459 ATGCTGTGTGTAATTCGTTCCAGGCTTATATACCTGTGTGACATGTTCCGTAGGGCTGATTT 2518
Qy 293 AapLeuAlaIleAlaArgAsnThrLeuIleGlyValAspAspPheLeuSerGlyGlnThr 312
Db 2519 GATCTGTGTAATAAAAAACAACCTGCTGATTTGGTGGAGAGACTGCAAGAGTGAGACAGACC 2578
Qy 313 LysMetLeuSerValLeuValAspPheLeuValGlyAlaGlyIleLys----- 328
Db 2579 AAGATGAATCTGTCTGTGTCATTTCTTTGGTGGTGGAAAT-AAAGGTGGAAACCTA 2637
Qy 328 ----- 328
Db 2638 GTATCTCTCTCTATTAAGATGAAGTATTTTGGCAATGACGTTATTGCAATTAATC 2697
Qy 329 -----ProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 2698 TTCTATATTTTCAATTTTCATGACAGCCCACTCAATCTGAGCTACACCACTTGAGGAAAC 2757
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGlyLeuSerLysSer 360
Db 2758 AACGATGGCATGACCTGCTCTGCCCTTCAACATTCAGTCCAGAGAGATCTCCAGAGAC 2817
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnHisPro 380
Db 2818 AACGTGTGATGACATGCTGTGAGCAATGACCATCTCTATAGAGCCCGGAGACATCCCC 2877
Qy 381 AspHisValValValIleLys----- 387
Db 2878 GATCATGTGTTGTCTCATYCAAGGTCTGTAGCTGATCTTCACTCGTTAAAGTTGACAT 2937
Qy 388 -----TyrValProTyrVal 392
Db 2938 ATGCAGAGCATTTTACATTGAAACTTGTCACTCTTTGTGACAGTATGTGCGTATCGTG 2997
Qy 393 GlyAspSerLysArgAlaMetAspGlyTyrThrSerGlnIlePheMetGlyGlyLysSer 412
Db 2998 GGAAGACAGCAAGAGGCGCTATGACGAGTACACCTCAGAGATCTTCAATGGCGGCAAGAAC 3057
Qy 413 ThrIleValLeuHisAsnThrCysGluAspSerLeuValAlaAlaProIleIleLeuAsp 432
Db 3058 ACCATGTGCTGACAAACACTGTGTGAGACTGCTGCTGCGCACCTTATCATCTTGAT 3117
Qy 433 LeuValLeuValAlaGluLeuSerThrArgIleGluPheLysAlaGluAsnGlu----- 450
Db 3118 CTGGTCTCTTGGCTGAGCTCAGACCAAGATCCACTGAAGCTGAGGAGAGATGAAGA 3177
Qy 450 ----- 450
Db 3178 GCCCCCAAGTATTAACCTGAAGAACGCTGACAGCTAGTATATAGCACTTTAATA 3237
Qy 451 -----GlyLysPheHisSerPheHisProValAlaThrIle 462
Db 3238 CCTTCTGTGTCTCTCTTATGACAGCAAAATTCACCTTCCACCCGGTGGCCACACATC 3297
Qy 463 LeuSerTyrLeuThrIleLysAlaProLeu----- 471
Db 3298 CTGAGCTTACTCACCAGGACCCCTGTGTAAGCTTTTCTCTGCATCCCGGATATCAG 3357
Qy 472 -----Val 472
Db 3358 CACTGGCTTTTGCTTCAATCCAGCACTGATGCTCTCTGAAACCTGAAACAAGGTT 3417
Qy 473 ProProGlyThrProValValAlaAlaLeuSerLysGlnArgAlaMetLeuGluAsnIle 492
Db 3418 CCCCCTGGACACCGGTGTGTGAACGCTTGGCCAAAGCAGACGCGCATGCTTGGAGAACATC 3477

```

```

Qy 493 MetArgAlaCysValGlyLeuAlaProGlnAsnAsnMetIleLeuGlyTyrLys 510
Db 3478 ATGAGGCGCTGCGTGGCTGCGCCCGAAGAAACAATCATCTGAGATACAG 3531

RESULT 5
US-09-677-064-15
/ Sequence 15, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ TITLE OF INVENTION: Plants and Uses Thereof
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677,064
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-677-064-15

Alignment Scores:
Pred. No.: 2,716-217 Length: 3546
Score: 1813.00 Matches: 443
Percent Similarity: 40.86% Conservative: 31
Best Local Similarity: 38.19% Mismatches: 36
Query Match: 68.88% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-2 (1-510) x US-09-677-064-15 (1-3546)
Qy 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
Db 60 ATGTTTCATCGAGAGCTTCGCGGTGAGAGCCCGACGTCGCGTACGCGCCGACGAGATC 119
Qy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 120 GAGTCGAGATACCGGTGTACGACACGAGAGCTGTGTACGAGGGGCAAGAACGCGCTCA 179
Qy 41 GlnTPrIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db 180 CGCTGGTGTCTCGCCCAAGTCTGTCAAGTACCACTTCGAGCAAGAACGCGCTCCCC 239
Qy 61 LysLeuGly----- 63
Db 240 AAGCTCGGGTATGATGAGATGACAGGCGCTAGCTCACTCTGTGAACCTCTCTCTC 299
Qy 63 ----- 63
Db 300 CCGTCTCAGTCAATCTCGTCGAGATCAACTGTGTGCGGTTCCTTAATCTTA 359
Qy 63 ----- 63
Db 360 TGAATACTTACTGCTTGTGCTGAAGACGAACGCTGTATTTGTGACAGCTACGACAC 419
Qy 63 ----- 63
Db 420 ACTTGCCCATCCGAGTGGGTCAATCAGCTGATTTGAATTGATTGATGTGCTT 479
Qy 63 ----- 63

```

Db 480 TTCCATATTTCATCATCTTGCCTACTCTGTGCATGATTACAGAAAGCTCCCTTCC 539
Qy 63 ----- 63
Db 540 TGAATCTGTCTAGGCTTTTGTCTGTGCAGCGTAGCTGATCAATTTGTTCA 599
Qy 63 ----- 63
Db 600 AGATCAAAATTCACAGGAGCAGTAGACAGACAGAACTCATTAACGTAGCAAAATGA 659
Qy 63 ----- 63
Db 660 TACTAGATTACTGCAAGTGTGCATACGGCGCAATCTGCATCTGACCCCTTTGTT 719
Qy 63 ----- 63
Db 720 TAATTCCTGTCTTANGATGTTGCTACGTGACGCTGTTGTGTATGTGTGACGC 779
Qy 63 ----- 63
Db 780 TGTCAAGCCGCTGTCTCTGTCTGACGAGATGCCAACTTTCTGTCTGTGTGTCAG 839
Qy 64 ValMetLeuValGlyTyrPglYglYAsnAsnGlySerThrLeuThrGlyValIleAla 83
Db 840 GTGATGCTTGTGGGTGGGGAGGCAACACGGSTCCACGTGACGGCTGGGCTCATTTGC 899
Qy 84 AsnArgGlu----- 86
Db 900 AGCAGGAGTGAAGTACTTAATTTGCTTATATGCTTTCGTTGTTTTCAGTTATTA 959
Qy 87 ----- 92
Db 960 ATGGCCTGACAGAACTGAATTTTGTGTGGCTGTTTCAGGGAGATCTCATGGCCGAC 1019
Qy 93 LysAspLysIleGlnGlnAlaAsnTyrPheGlySerLeuThrGlnAlaSerAlaIleArg 112
Db 1020 AAGGACAAAGGTGACAGCAAGCACTACAGGCTC-CTCACCAAGGCTTCCACATCAAG 1078
Qy 113 ValGlySerPheGlnGlyGlnIleTyrAlaProPheLysSerLeuProMet--- 131
Db 1079 GTGCGAGCTACAAAGGGAGAGATCTATGCGCTTCAAGAGCTCTTCCCATGTGA 1138
Qy 131 ----- 131
Db 1139 ACTATTAATAGACTGACTAATATCTCTTTTACTGAAACCAACATACATACAAAG 1198
Qy 131 ----- 131
Db 1199 CATATTCGGTAAGTGTCTAGTTGATGTATTAAGTGAACCTGTCTTCAAGCCAGTGT 1258
Qy 131 ----- 131
Db 1259 TCNAAGTAAACGAATGTATCATGTGGTTGAAAAAACAAGGTCTAAATTTGTGAAG 1318
Qy 131 ----- 131
Db 1319 GAATGTTAACTTAGCATATGAAAGGGAGACATTTGAAGAGTGTGTAACATC 1378
Qy 131 ----- 131
Db 1379 GACTCGGTGCGCATGTGTGATCTTAAGCTTGAAGTCAAAAGGTCTGTGGTTAGCTCAA 1438
Qy 131 ----- 131
Db 1439 AGGTTCCAGCATTTGATTAACATCTGTGCAATATTGAAGATGTGTAACGTGTTGA 1498
Qy 131 ----- 131
Db 1499 GATTGCGACATCTCGGTGTTGTATCTCTTTTCATGACCAAGTGTGAACGTGTTTGA 1558
Qy 132 -ValAsnProAspAspIleValPheGlyGlyTyrPheIleSerAsnMetAsnLeuAla 151
|||||

Db 1559 GGTGAACCCAGACGACATTTGTGTCGAGGCTGGGACATTAGCAACATGACCTGGCCGA 1618
Qy 151 PalAsnMetAlaArgAlaLysValPheAspIleAsnLeuGlnLysGlnLeuArgProTyrIle 171
Db 1619 CTCCATGTACAGGCGCCAAAGTGTCTGATATTTGACCTGCAGAAAGCAGCTCAGCCCTACAT 1678
Qy 171 TglLysMetLeuProLeu-ProGlyIleTyrAspProAspPheIleAlaAsnGlnG 191
Db 1679 GAGTTCATGTGTGCCACTTCCCGGATATCTATGATCCGACTTTCATCGCGCTTAACAG 1738
Qy 191 LngLysArgAlaAsnAsnValIleLysGlyThrLysGlnGlnValGlnGlnIleIle 211
Db 1739 GCTCTGCGGCCAACAGTGTCAACAGGCGCACCAAGAAAGAACAGGTCAGATCATCA 1798
Qy 211 LysAspIle----- 213
Db 1799 AGGATATCAGGTATATGATATGATGATGCTTAAGCTGTGTGCTAAGGTGACCCAGTG 1858
Qy 213 ----- 213
Db 1859 CAACCTAAACAAATTAATTAATCTAATGAATTTGGTAATTAATACATATACAGAACAT 1918
Qy 213 ----- 213
Db 1919 ATTTGTTAACGGGTGACCTTAGAAGTGTGCATGTATGTGACAAATTTGACATTCGATA 1978
Qy 213 ----- 213
Db 1979 TACAGTACCGCTCACTTGCAATGAGGACTCACAAAGAACTAAACTAGAAAGCTTAA 2038
Qy 213 ----- 213
Db 2039 GCAACTATTCGTAGCTAATGATGATTTGTGTGACATGTTGAAGATCTAATTAACGT 2098
Qy 213 ----- 213
Db 2099 GGTGAAGAAATATGTTCACTAGCAATAAGTATATCCATACAGAAAGCTATGATA 2158
Qy 214 -----LysAlaPheLysGlnAlaThrLysValAspLysValValLeuTyrT 230
Db 2159 GCTAATGAACACAGGAGTTTAAGAGAAAGAACAAAGTGACAAAGTATGTTGTGTGGA 2218
Qy 230 hrAlaAsnThrGlnArgTyrSerAsnLeuValValGlyLeuAsnAspThrMetGluAsn 250
Db 2219 CTGCAACACCTAAGAGTATACCAATGTGTGCGCTGTCTCAACGACACAGTGAAGATC 2278
Qy 250 eulLeuAlaIleValAspArgAsnGlnAlaGlnIleSerProSerThrLeuTyrAlaIleA 270
Db 2279 TACTGGCATCTGTGACAAAGACGAGCGAGATTCACCATCAACATATATGCAATG 2338
Qy 270 LysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro 289
Db 2339 CCTGTGTCAATGAGGGGCGCGTTCATCAATGAGGAGCCCCAGAACCTTTGTGCTGT 2398
Qy 289 ----- 289
Db 2399 GTGCGGTGTTGTGTGTGTTGCAAAAGCTTCAAGGTGTGCAATTCGTGCCAAAGTTTC 2458
Qy 290 -----GlyLeuIle 292
Db 2459 ATGTGTGCTATTTCCGTTCCAAAGCTTATTAATCTGTGATGTGCTGAGGCTGANT 2518
Qy 293 AspLeuAlaIleAlaArgAsnThrLeuIleGlyIleAspAspPheLysSerGlyGlnThr 312
Db 2519 GATCTTGCTATTAACAACTGCTGATGTGTGTGACGACTTCAAGATGAGACAGACC 2578
Qy 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLys----- 328
Db 2579 AAGATGAATCTGTCTGTGTGATTTCTTGTGTGTGCTGAAT-AAAGGTGGAACTTA 2637
Qy 328 ----- 328
Db 2638 GTATCTCTCTTCTATTAAGATGATGTTTGTGGCAATGACGTTATGTGCAATATCTC 2697

QY	329	-----	ProthrSer11LeuValSerTyrAsnHisLeuGlyAsn	340
Db	2698	TTCTAATATTTTCATTTTCATGACGCCCACTTCATATCGAGCTCAACACCACTTGGAAC	2757	
QY	341	AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer	360	
Db	2758	AACGATGGCATGAACCTGTCTCGCCCTTCAACATTCAGCTCCAAAGAGATCTCCAAAGC	2817	
QY	361	AsnValValAspAspMetValAsnSerAsnAlaIleuTyrGluProGlyGluHisPro	380	
Db	2818	AACGTGGTGATGACATGCTCTCGAGCAATCCATCTCTATGAGCCCGGAGCATCTCC	2877	
QY	381	AspHisValValValIleLys-----	387	
Db	2878	GATCATGTCGTGTTCATCAACAGTCTGTAGCTGATCTTTACCTCGTTAAAGTTGACAT	2937	
QY	388	-----	TyrValProTyrVal	392
Db	2938	ATCGAAGCGAGATTACATTCGAAACTTGTCTACTCTTTGTGTGCAGTATGTGCTGACTG	2997	
QY	393	GlyAspSerLysArgAlaMetAspGluTyrThrSerGluIlePheMetGlyLysSer	412	
Db	2998	GGAAGACAGCAAGAGGGCTATGACAGATACACTCAAGATCTTCATGCGCGCAAGAAC	3057	
QY	413	ThrIleValLeuHisAsnThrCysGluAspSerLeuMetAlaIleProIleIleLeuAsp	432	
Db	3058	ACCATTCGTGCGCAACACACTGTGAGGACCTGCTCTCGCGGACCATATATCTTGAT	3117	
QY	433	LeuValIleuLeuAlaGluLeuSerThrArgIleGluPheLysAlaGluAsnGlu-----	450	
Db	3118	CTGATGCTCTTGCGTGAAGCTAGCACCAGATTCACGCTGAAGCTGAAGGAGAGTAAGA	3177	
QY	450	-----	450	
Db	3178	GCCTCCCAAGATTAACCTGAAGACGCTGACGCTAGGTATATAGCACTTTTAATA	3237	
QY	451	-----	GlyLysPheHisSerPheHisProValAlaThrIle	462
Db	3238	CCTTCTGGTGTCTCTCTTATGACAGACAAATTCACACTCTTCCACCGGTGGCCACATC	3297	
QY	463	LeuSerTyrLeuThrLysAlaProLeu-----	471	
Db	3298	CTGAGCTACCTCACCAAGGACCCCTGGTAAGCTTTTCTCTGCATCCGGGATCACTG	3357	
QY	472	-----	Val	472
Db	3358	CACGCGCTTTGCTTCATTCAGCAACATGATGCTCTCTGAAACCTGAACAACAGGTT	3417	
QY	473	ProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIle	492	
Db	3418	CCCCCTGGCAACACCGGTGTGAACGCTCTGCGCAAGAGAGCGGAGTGTCTGAGAAACATC	3477	
QY	493	MetArgAlaCysValIleGlyLeuAlaProGluAsnAsnMetIleLeuGluTyrLys	510	
Db	3478	ATGAGGGCTCTGCTTGCGCTGGCCCCCAAGAAACAATGATCTTGAAGTACAG	3531	
RESULT 6				
US-09-118-442-14				
Sequence 14, Application US/09118442B				
Patent No. 6197561				
GENERAL INFORMATION:				
APPLICANT: Martino-Calt, Susan J.				
APPLICANT: Wang, Hongyu				
APPLICANT: Beach, Larry R.				
APPLICANT: Wang, Xun				
APPLICANT: Bowen, Benjamin A.				
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in				
FILE REFERENCE: 0706				
CURRENT APPLICATION NUMBER: US/09/118,442B				
EARLIER APPLICATION NUMBER: 60/055,446				

```

: EARLIER FILING DATE: 1997-08-11
: EARLIER APPLICATION NUMBER: 60/055, 526
: EARLIER FILING DATE: 1997-08-08
: EARLIER APPLICATION NUMBER: 60/053, 944
: EARLIER FILING DATE: 1997-07-28
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 14
: LENGTH: 3546
: TYPE: DNA
: ORGANISM: Zea mays
: OS-09-118-442-14

```

Alignment Scores:

Pred. No.:	3,62e-217	Length:	354
Score:	182.	Matches:	442
Percent Similarity:	40.86	Conservative:	32
Best Local Similarity:	38.10	Mismatches:	36
Query Match:	68.84	Indels:	652
DB:	3	Gaps:	9

US-10-718-952-2 (1-510) X US-09-118-442-14 (1-3546)

OY		1 MetPheIIeGluAAsPhelyValGIcyProAsnValLysrYrrThGluThrGluIle	20
Dd	60	ATGTTCATTGAAGCGCTTCGGCTGCAGAGACCCCAAGTGGCGTAACGGCCGAGAAATC	119
OY	21	GlnSerValTyraSnTyrgIuThrThrgIueValInIsGluAmArGsnGlyThrTyr	40
Dd	120	GAATCCGAGTAGCCGGTACACACAGCGAGCTGGTACACGAGGGCACAGAACGGCGCTCA	179
OY	41	GlnTrpIleValIyAsProLySserValIySrTyrgIuPhelysrThraenIleHisValPro	60
Dd	180	CAGTCGGATGCGTCCGCCCAAGTCCCTCAAGTACAATTCCGACCAGAACCGCGCTCCC	239
OY	61	LysLeuGly-----	63
Dd	240	AAGCTGGGATATGATACGATGACGAGCGCCCTAGCTCATCTCTGTGTAAACCTCTCTC	299
OY	63	-----	63
Dd	300	CCGTGCTCAgTCAAATCTCCGTGCAGATCAAAGTGtGGGGTTCCCTTAATCTTA	355
OY	63	-----	63
Dd	360	TGAAAATCTTACTGCTTTTGTCTGAGAAGAACCGTCAATTGTGTACAGCTACGACAC	419
OY	63	-----	63
Dd	420	ACTTGCCATCCGGATGCGTCAAATCAGCTCGATTGAATTCGATTCAGTGtGCCCTT	479
OY	63	-----	63
Dd	480	TTTCATATTTCATCATCCCTGCGCTACTGTGCATGTATACAGAAAAGTCTTTCTC	539
OY	63	-----	63
Dd	540	TGAACTTGTCTTAGGCTTTTGTCTGTGCACGTAGCTGTATCAATTTGTTCATGTA	599
OY	63	-----	63
Dd	600	AGATCAAAATTCAGACGAGGACGATGAGACAGACAGAACTCAATTACATAAGCAAATTGA	655
OY	63	-----	63
Dd	660	TACTAGAGATTACGCAGATGTGCATACGGCGCAATCTGCATCTTGACCCCTTTGTT	719
OY	63	-----	63
Dd	720	TAATTCCTGTTCATGATGTTCGCTACGTGCAGCTGTTGTGTATGATGTGACAGC	779
OY	63	-----	63

Db 780 TGTCAACCCGCTGTCTCTGTCCGACGAGANTGCCAACTTTTCTGTCTGTGTGCAGG 839
QY 64 ValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyValIleAla 83
Db 840 GTGATGCTGTGGGGTGGGAGGCAACAAGGCTCCACGCTGACGGCTGGGCTCATGTGC 899
QY 84 AsnArgGlu----- 86
Db 900 AACAGGAGTGAAGTACTACTAATTGTCTATATTGCTTCCGTTGTTTCAGTTATTA 959
QY 87 -----GlyIleSerTyrAlaThr 92
Db 960 ATGGCTTAACAGAACTGAATTTGTGTGTGTGTTCAGGGAGATCTCATGTGCCGACC 1019
QY 93 LysAspLysIleGlnGlnAlaSerTyrPheGlySerLeuThrGlnAlaSerAlaIleArg 112
Db 1020 AAGGACAAAGGTGCAGACAACTACTACGGCTC-CTCAACCGAGCTCCACCATCAGA 1078
QY 113 ValGlySerPheGlnGlyGluGluIleTyrAlaProPheLysSerLeuProMet----- 131
Db 1079 GTGGGAGCTACAAAGGGAGAGATCTATGCGCGTTCAAGAGCCTCTCCCATGGTA 1138
QY 131 ----- 131
Db 1139 ATCTATTATAGACTTGACTAATACTCTCTTTTACTGAAACCAACATACATACAAAG 1198
QY 131 ----- 131
Db 1199 CATATTCCTGAAGGTGTAGTTGANTGTATAAATGAACCTGTCTTCAGGCCAGTGTCTC 1258
QY 131 ----- 131
Db 1259 TCAAGTAAACGGAATGTATATCTATGGGTGTAAGAAAAACAAGTCTAATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAAAGTTAACTTAGCATATATGAAAGGGAGAGACTGTAAAGAAAGTGTGTAACAATC 1378
QY 131 ----- 131
Db 1399 GACTCGGTCTGCATGTGTGTGATCTCTACTTGCAAGTCAAAAGTCTGTGTAGCCAA 1438
QY 131 ----- 131
Db 1439 AGGTTCACAGCATTTGATTCACACTGTGCAATTAAGATGGTCTTAACCTGGTGA 1498
QY 131 ----- 131
Db 1499 GATTGCGAGACTCGGTGTGTGTATCTTTTCATGACCAAGTGTAAACGTGTTTCA 1558
QY 132 -ValAsnProAspAspIleValPheGlyGlyTyrAspIleSerAsnMetAsnLeuAlaAs 151
Db 1559 GGTGAACCCAGACGACATGTGTGTGAGGCTGGGACATTACCAACATGAACCTGGCCGA 1618
QY 151 PALAMEALAAAGAlaLysValPheAspIleAspLeuGlnLysGlnLeuArgProTyrPhe 171
Db 1619 CTCCTAGACCAAGGCGCAAGCTCTGATATTAACCTGCAGAGACGCTCAAGCCCTACAT 1678
QY 171 TGIuSerMetLeuProLeu-ProGlyIleTyrAspProAspPheIleAlaIleAsnGln 191
Db 1679 GGAAGTCATGGTGCACCTTCCCGGATCTATGATCCGACCTTCACTGGGCTTAACAGG 1738
QY 191 LUGluArgAlaAsnAsnValIleLysGlyTThrLysGlnGlnGlnValGlnGlnIleIleL 211
Db 1739 GCTCTCGCGCCAAAGTGTCAATCAAGGCGCAACCAAGAAAGAACAGGTGAGCATCATCA 1798
QY 211 LysAspIle----- 213
Db 1799 AGGATATACGGTATATGATATGATATGATCTTAACGTGCTTGGTCTAAGGTGACCAAGT 1858
QY 213 ----- 213
Db 1859 CAACCTAAACAAATAAATACTACTATGAATTTGGTAAATATACATACATACAGAGCAT 1918

QY 213 ----- 213
Db 1919 ATTGTTTAAACGGTGCATTTAGAGAGTCTGCATGTATGTGTGACAAATTTGACATTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGCTCACTTCGATGAGAGCTCCAAAGAACTAAACTACTGAAGAGTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTCGTAGCTAATGATGATATTGTGTGACATGTGTTGAAGATCTAAGTTAAGT 2098
QY 213 ----- 213
Db 2099 GGTGAAGAAATATGTTCACTAGTATTAAGTATTCATTACAGAACATGGCTTATGTA 2158
QY 214 -----LysAlaPheLysGluAlaThrLysValAspLysValValIleuTyrT 230
Db 2159 GCTAATGAACACGGAGGAGTTTAGGAGACAGAACAAAGTGAACAAAGATAGTTGTGTGTGA 2218
QY 230 hrAlaAsnThrGluArgTyrSerAsnLeuValGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACACTGAAGGTATAGCAATGTGTGCGCTGTCTCAACGACAGATGAGAAATC 2278
QY 250 euLeuAlaAlaValAspArgAsnGluAlaGluIleSerProSerThrLeuTyrAlaIleA 270
Db 2279 TACTGGCATCTGTGACAAAGACGAGGCGAGGTATCACCATTCACATATATGTCATTTG 2338
QY 270 LAcYeuValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro- 289
Db 2339 CCTGTCTATGAGAGGGGGTCCGTTTCATCAATGGAGCCCAAGAACACTTTGTGCTCTG 2398
QY 289 ----- 289
Db 2399 GTGCGGTGTGTGTGTGTGCAAAAGCCTCATGGNGTTGCATTTCTGTCCAAAGTTTC 2458
QY 290 -----GlyLeuIle 292
Db 2459 ATGGTGTGTATTTCTGTCCAGGCTTATATACCTGTTCAGTTCGTAGGGCTGATTT 2518
QY 299 AspLeuAlaIleAlaArgAsnThrLeuIleGlyGlyAspAspPheLysSerGlyGlnThr 312
Db 2519 GATCTTGCTATTAATAAACACTGCTTGATGTGTGTGAGCATTCGAAGTGTGACAGACC 2578
QY 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLys----- 328
Db 2579 AAGATGAATACTGTCTGTGTGATTTCTTGTGTGTGTGGAAT-AAAGGTGGAACCTA 2637
QY 328 ----- 328
Db 2638 GTATCTCTTCTATTAAAGATGAAGTGTGTTTGTGGCAATGACGTATATGCAATAACTC 2697
QY 329 -----ProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
Db 2698 TTCTATATTTTCATTTTCATGACGCCACCTCAATCGTAGCTACAAACACTTGGGAAAC 2757
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysSer 360
Db 2758 AACGATGCGATGAACCTGTCTGCCCTTCAACATTCAGGTCCAAAGAAATCTCCAAAGAC 2817
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlyIlePro 380
Db 2818 AACGTGTGTGATGACATGTGTCTGAGCAATGCAATCTCTATGAGCCCGGCGAGCATCC 2877
QY 381 AspHisValValValIleLys----- 387
Db 2878 GATCATGTGTGTGTATCAAGGTCTGTAGCTGATCTTTCACTGTGTTAAAGTTGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGGACATTTACATTTGAACCTGTGCACTCTTTTGTGACGATATGTGCGTACGG 2997

QY 393 GLYASPERLYSARGALANERASPGIUTYRTHRSERGIULLEPHMEGLYLILYSER 412
DB 2998 GGAGACAGACAGAGGCTATGACGAGTACCTCAGAGATCTTCATGGCGCAAGAAC 3057
QY 413 THRILEVALLEHIASENTHRCYSGIUAAPERSELEULEVALIALAProIlelleuuep 432
DB 3058 ACCATCGTCTGCACAAACCTGTGAGACTGTGCGCGCACATCAATCTTGAT 3117
QY 433 LEUVALLEULEVALIGIUEUSERTHRARGIILEGIUPHELYSALIGIUBN----- 449
DB 3118 CTGGTCTCTTGGCTGAGTCCAGCACGAGATCCAGCTAAAGCTGAGGAGGGTAA 3177
QY 449 ----- 449
DB 3178 GCGCCCAAGTATTAACCTGAAGACGCTGACGCTAGTATATACACTTTAATA 3237
QY 450 -----GIUGLYLSPHEHIASEPHENIProVALIAthrlle 462
DB 3238 CCTTGTGTGTCTCTTATGAGAGCAAAATTCACCTTCCACCCGGTGGCCACATC 3297
QY 463 LEUSERLYLEUTHRLYSALAPROLEU----- 471
DB 3298 CTGAGCTACTCAACAGGACCCCTGTAAAGCTTTCTCTGCAATCCGGCATCATC 3357
QY 472 -----VAL 472
DB 3358 CACTGGCTTTGCTTCAATCCAGCAGCATGATGCTCTCTTGAAACCTGAACAGAGTT 3417
QY 473 PROPROGLYTHRProVALIaenIALeUSERLYSGLIARGLAMetleuIuueIle 492
DB 3418 CCCCCGACACACCGGTGTGAACGCTGCGCAACAGACGCGATGCTGGAAGAACATC 3477
QY 493 METARGALACYSVALIGIUEUALAPROGIUASAMetIleleuGIUTYRILYS 510
DB 3478 ATGAGGCGCTGCGTGGCTGGCGCCCAAGAACATATCTCTGAGTRCAAG 3531

RESULT 7
US-09-677-064-14
; Sequence 14, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Calc, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-14

Alignment Scores:
Pred. No.: 3,62e-217 Length: 3546
Score: 1812.00 Matches: 442
Percent Similarity: 40.86% Conservative: 32
Best Local Similarity: 38.10% Mismatches: 36
Query Match: 68.84% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-2 (1-510) x US-09-677-064-14 (1-3546)
QY 1 MetPheIleGIUaenPheLYSVALIGIUCYSPROaenVALYRTHRGluThRGluIle 20
DB 60 ATGTTCAATCGAGAGCTTCGGGTGAGAGCCGCCACGTCGGGTACAGGCCGACGAGATC 119
QY 21 GINSERVALYRAsnTYRGLIUTHRTHRGluLeuVALHISGLIaenARGAsnGIYthRYR 40
DB 120 GAGTCGAGATACCGGTACGACACGACGAGCTGTACACGAGGCGAAGACGCGCCCTCA 179
QY 41 GINTPILIEVALYSPROLYSERYALYSTRYRGIUPHELYSTHRAmIleHISVALPro 60
DB 180 CGCTGGGTGTCGCCCAAGTCGTCAAGTACAACTTCGAGACCAAGCCGCTCCCC 239
QY 61 LYSLEUGLY----- 63
DB 240 AAGCTCGGGTATATGACGATGACGGGCCCTAGCCTCATCTGTGAACCTCTCTCTC 299
QY 63 ----- 63
DB 300 CCGTCTCAGTCAAACTCTCCGTGAGATCAACTGTGCGCGTTCCTCTAATCTTAA 359
QY 63 ----- 63
DB 360 TGAATACTTACTGCTTGGCTGAAAGACGAACCGTGTAAATTGTTGACAGCTACGACAC 419
QY 63 ----- 63
DB 420 ACTTGCCATCCGAGATGGGTCAATCAGCTGATTGAAATTCATCGATGATGTCCTT 479
QY 63 ----- 63
DB 480 TTCATATTTGCATCATCCTCGCTACTGTCATGATTAACAGAAAGTCTTTCTCTC 539
QY 63 ----- 63
DB 540 TGACTTGTCTTAGGCTTTTGTCTGTGACGTGAGCTGTATCAATTTGTATATGA 599
QY 63 ----- 63
DB 600 AGATCAATTCAGACAGGAGAGATGACGACGACAGACAACTATTACACTACAAATTGA 659
QY 63 ----- 63
DB 660 TACTAGATTACTGGCAAGTGTGATAGCGCGCAATCTGCCATCTCGAACCCCTTTGTT 719
QY 63 ----- 63
DB 720 TAATTCCTGTTCTATGCAATGTTGCCATGACGTGACGCTGTGTGTATGATGTCAGGC 779
QY 63 ----- 63
DB 780 TGTCAACCGCTTGTCTGTGTCCAGAGATGATGCCAATTTTCTGTCTGTGTGTCAGG 839
QY 64 VALMetLeuVALIGIYTRPGIYGLIaenANGIYSERTHRLeuTHRGlyVALIleIle 83
DB 840 GTGATGCTTGTGGGTGTGGGAGCAACACGGGTCCAGCTGACGCGTGGGTCTATTGCC 899
QY 84 AsnARGIu----- 86
DB 900 AACAGGAGGTGAGTACTTAATTGTCCTATATTGCTTTCCTTCCTGTTTTCAGTTATTA 959
QY 87 -----GLYIleSTHRAmIle 92
DB 960 ATGCGCTAACAGAACTGAATTTTGTGTGTGTTGTTTCAGGGATCTCATGGCGAC 1019
QY 93 LYSAspLYSILEGLINIAAsnTYRPhEGIYSERLeuTHRGluAlaSERAlaIleARG 112
DB 1020 AAGGACAGAGTGCACAGCAACTACTACGGCTC-CTCACCCAGGCGCTCCACATCAGA 1078
QY 113 VALIGIYSERPhEGINGLIGIUGIUIleTYRAlaProPheLYSERYLeuSERMet--- 131
DB 1079 GTTCGAGCTCAACGCGGAGAGATCTATGCGCGCTTCAGAGCCTCTTCCATGTGA 1138

QY 131 ----- 131
Db 1139 ATCTATTATAGACTTGACTTAATCTCTTTCTTTTACTGAAACCAATACATATACAAAG 1198
QY 131 ----- 131
Db 1199 CATATTCCGTAAAGTCTAGTTATGTATATAATGAACCTGTCTTTAGCCAGTGTCTC 1258
QY 131 ----- 131
Db 1259 TCAGATAAAGCATGTATATCATTTGGGTTGAAAAAACAAGGTTCTAATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAAAGTTAAACTTAGATATGAAAGGGAGACCTGTAGAAGGCTGTGAACAATC 1378
QY 131 ----- 131
Db 1379 GACTGGGTCTGCATGTGTGATCTTACCTTGCAAGTCAAAAAGTCTGTGTGTTAGCCAA 1438
QY 131 ----- 131
Db 1439 AGGTTCCAGCATCTTTGATTAACCTGTGAGATTGACGATGTGCTAACTGTGTGCA 1498
QY 131 ----- 131
Db 1499 GATTCGACAGCTCGGTGTTTATCTCTTTTCATGACCAAGTGTAACTGGTTTCA 1558
QY 132 -ValAenProAspAsp11eValPheGlyGlyTrpAsp11eSerAsnMetAsnLeu11aAs 151
Db 1559 GGTGAACCCAGACGACATGTGTTCGAGGCTGGGACATTTAGACATGAACTGGCCGA 1618
QY 151 pAlaMetAlaArgAlaLeuValPheAsp11eAspLeuGlyGlyLeuAspProTyrMe 171
Db 1619 CTCATGACACAGGGCCAGAGTGTGATATGACCTGACAGAACGACTGAGCCCTACAT 1678
QY 171 LeuSerMetLeuProLeu-ProGly11eTyrAspProAspPhe11eAla1aAsnGln 191
Db 1679 GAGATCCATGATGTGCCCTTCCCGGATCTATGATCCGACTTCATCGCGCTAACCAAG 1738
QY 191 LuGluArg11aAsnAsnVal11eLeuGlyThrLeuGlnGluGlnGln11e11eLeu 211
Db 1739 GCTCTGCGCCACAGTGTCAATCAAGGGCACAAAGAAAGACAGTGAAGCATCATCA 1798
QY 211 yAsp11e----- 213
Db 1799 AGGATATCAAGTATATGATGATGATGATGCTAACGTGCTGTGCTAAAGTGACCCAGTG 1858
QY 213 ----- 213
Db 1859 CAACCTMAAACAAATAAATACTACTATGATTTGTTAAATATACATATCATCAGAGCAT 1918
QY 213 ----- 213
Db 1919 ATTGTTTAAACCGGTGACTTAGAGTCTGCATGCTATGTTGACAAATTTGACATTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGGCTCACTTGATGAGAGCTCCACAAAGAACTAAACTAGTGAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTCTAGTATATGATGATGATTTGGTGACATGTTGAAAGTCTAGATTAACTG 2098
QY 213 ----- 213
Db 2099 GGTGAGAAATATGTGTCACTAGTATTAAGTAATCCATTACAGAACAAATGCTTATGTA 2158
QY 214 -----LysAlaPheLeuGlyAlaThrLeuValAspLeuValValLeuTyrP 230
Db 2159 GCATATGAAACAGGAGTTTAGGAGAGAAAGAACAAAGTGGACAAAGATAGTTGTGTGGA 2218

QY 230 hrAlaAsnThrGluArgTyrSerAsnLeuVal1aGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCACACACCTGAAGAGATATAGCATGTGTGGCTGTGTCTCAACGACAGATGAGAAATC 2278
QY 250 euLeuAlaAlaValAspArgAsnGluAlaGly11eSerProSerThrLeuTyrAla11eA 270
Db 2279 TACTGCATCTGTGGACAAAGAACGAGCGGAGGATCACATCAACATATATAGCCATTG 2338
QY 270 lAcysValMetGluAsnValProPhe11eAsnGlySerProGlnAsnThrPheValPro 289
Db 2339 CCTGTCTATGAGGGGGTGGCCGTTCATCAATGGGAGCCGCCAGAACACCTTTGTCTGT 2398
QY 289 ----- 289
Db 2399 GTGCGTGTGTGGTGTGTTGCAAAAGCCTCATGTGTTCATTTCTGTCCAAAGTTTC 2458
QY 290 -----GlyLeu11e 292
Db 2459 ATGGTGTGATTTCTGTTCCAAAGCTTATTAACCTGTGACATGTTGATGGGCTGATT 2518
QY 293 AspLeuAla11eAlaArgAsnThrLeu11eGly11aAspAspPheLeuSerGly11eThr 312
Db 2519 GATCTTGCTATAAAAACAACCTGCTGATTTGGTGTGACGACTTCAAGAGTGAACAGACC 2578
QY 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGly11eLys----- 328
Db 2579 AAGATGAAATCTGTCTGTGATTCATTTCTGTGGTGTGGAAT-AAAGTGGAAACTTA 2637
QY 328 ----- 328
Db 2638 GATATCTCTTCATTAAGATGAAGTTTTTTGGCAAAATGACGTATTTGCAATAATC 2697
QY 329 -----ProThrSer11eValSerTyrAsnH1sLeuGlyAsn 340
Db 2698 TTCTATATTTCATTTTCATGACGCCACCTCAATGTAGACTACAAACACTTGGGAAAC 2757
QY 341 AsnAspGlyMetAsnLeuSer11aProGlnThrPheArgSerLeuGly11eSerLysSer 360
Db 2758 AACGATGGCATGAACCTGTCTGCCCTTCAAACTTCAGTTCAAAGAGATCTCCAAAGAC 2817
QY 361 AsnValValAspAspMetValAsnSerAsnAla11eLeuTyrGluProGlyGluH1sPro 380
Db 2818 AACGTGTGTGATGACATGCTCTCGACCATTCGATCTCTATGAGCCCGGACAGCATCCC 2877
QY 381 AspHisValVal11eLys----- 387
Db 2878 GATCATGTCTGTGATCAAGGCTGTAGTATCTTCACTCGTTAAAGTTGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGGAGATTTACATTTGAAACTTGTCACTCTTTGTTGSCAGATATGCGGTAAGTG 2997
QY 393 GlyAspSerLysArgAlaMetAspGlyTyrThrSerGly11ePheMetGlyGlyLysSer 412
Db 2998 GAGACACGACAAAGGCTATGACAGATACACTCAGAAATCTTCAATGAGCGGCAAGAAC 3057
QY 413 Thr11eValLeuH1sAsnThrCysGluAspSerLeuLeuAla1aPro11eLeuAsp 432
Db 3058 ACCATGTGCTGCACAAACCTGTGAGACTGCTGCTCGCCGACACTATCATCTTGAT 3117
QY 433 LeuValLeuLeuAlaGlyLeuSerThrArg11eGluPheLysAlaGlyAsn----- 449
Db 3118 CTGTGTCTCTGTGCTGAGCTCAGACACAGATTCAGCTGAAGCTGAAGGAGGGGTAAAG 3177
QY 449 ----- 449
Db 3178 GCCCCCAAGATTTAATCTGAAGACGCTGACGCTAGGTATAGCACTTTAATA 3237
QY 450 -----GluGlyLysPheH1sSerPheH1sProVal1aThr11e 462
Db 3238 CCTTCTGTGTCTCTTATGAGGACAAATTCACCTCTTCAACCGGTGACACATC 3297
QY 463 LeuSerTyrLeuThrLeuValaProLeu----- 471

```
Db      3298 CTGAGCTACCTCACCAGGACCCCTGGTAAGCTTTCTCCGATCCGGACATCATCTG 3357
Qy      472 -----Val 472
Db      3358 CACTGCGTTTCTTCATCCAGCAGCATGATCTCTCTTGAACCTGACACAGAGTT 3417
Qy      473 ProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGlnAsnIle 492
Db      3418 CCCCCTGGGACACCGGTGGTGAACCGCTCTGGCCAGGACGAGCGGATGCTGGAGAACATC 3477
Qy      493 MetAlaGAlaCySValGlyLeuAlaProGlnAsnAsnMetIleLeuGlnTyrLys 510
Db      3478 ATGAGGCGCTGCGTTGGGCTGGCCCCAGAGAACAAATGATCTCGAGTCAAG 3531

RESULT 8
US-09-734-237B-72
/ Sequence 72, Application US/09734237B
/ Patent No. 6818752
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bui, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 72
/ LENGTH: 1602
/ TYPE: DNA
/ ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-72

Alignment Scores:
Pred. No.: 1,31e-163 Length: 1602
Score: 1382.50 Matches: 277
Percent Similarity: 69.27% Conservative: 86
Best Local Similarity: 52.86% Mismatches: 140
Query Match: 52.53% Indels: 21
Gaps: 8

US-10-718-952-2 (1-510) x US-09-734-237B-72 (1-1602)
Qy      3 IlegluAsnPhelYValGluCySProAsnValLysTyrThrGluThrGluIleGlnSer 22
Db      25 ATCACCCTCCGTTAAAGTACTTACCGACAGATGCACGTCACAGACAAACAGCTGCTCAC 84
Qy      23 ValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyrGlnTyr 42
Db      85 AAGTACAGCTACGAAATATCT---GTAGTTACGAAACAGCTAGTGGCCGCTTCGAT--- 138
Qy      43 IleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro---Lys 61
Db      139 ---GTAGAGCCACAGCTTCACAGCTACGCTTCAAACCTGACTGAAACAAACCGGAAAAA 195
Qy      62 LeuGlyValMetLeuValGlyTyrPglYglYAsnAsnGlySerThrLeuThrGlyVal 81
Db      196 CTAGGAATATATCTCTTGGGTGAGGTGGCAACAAAGGCTCCACCTTACGTGCGCTCGGTA 255
Qy      82 IleAlaAsnArgGluGlyLysSerTProAlaThrLysAspLysIleGlnGlnAlaSerTyr 101
Db      256 TTGGCGAATATAGACAAATGAGATTCAAACTACAGGAAAGCGCTTGAACCAACCACTAC 315
Qy      102 PheGlySerLeuThrGlnAlaSerAlaIleArgValGly---SerPheGlnGlyGlnGlu 120
Db      316 TTCGGCTCATATGACTCAATGTTCTACCTTGAAGAACTGGTATCGATCGGAGGGAATGAC 375
Qy      121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspIleValPheGly 140
Db      121 -----
```

```
Db      376 GTTATAGCTCTTTTAACCTCTCTGTGGCCATNGTTTACCCCAACGACTTTGTCTCT 435
Qy      141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db      436 GGTGGACATCATTAATACGACATGATATACGAAAGTATACGAGAGAAATGCAAGTTCTCGAA 495
Qy      161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyLys 180
Db      496 TATGATCTGCAACAAACCTTGAAGCGCAAGATGCTCTTGAGAACCTCTTCTTCATTT 555
Qy      181 TyrAspProAspPheIleAlaAlaAsnGlnGluAlaArgAlaAsnAsnValIle----- 198
Db      556 TACTAACCTGATTTATTTAGCTACCTATACAGATGAGAGCCATATGCTCATCAATTGG 615
Qy      199 -----LysGlyThr-----LysGlnGlnValGlnGlnIleLys 211
Db      616 GATGAAAAAGGCAACGTAAACCAAGAGGGGTAAAGTGAACCATCTGCAACGATCAGAGCG 675
Qy      212 AspIleLysAlaPheLysGluAlaThrLysValAspLysValValLeuThrThrAla 231
Db      676 GATATCCAGAAATTTCAAGAAAGAAACCGCTTGATTAAGTATGTTCTTGGACTGCA 735
Qy      232 AsnThrGluArgTyrSerAsnLeuValGlyLeuAsnAspThrMetGluAsnLeuLeu 251
Db      736 AATACTGAGAGGTACGTAAAGTATCTCTGCTGTGTAATGACCAATGAGAAACCTCTTG 795
Qy      252 AlaAlaValAspArgAsnGluAlaGluLysSerProSerThrLeuTyrAlaIleAlaCys 271
Db      796 CAGTCTATTAAGAAATGACATGAAGATGCTCTCTTCCAGATCTTGGACGACACATCT 855
Qy      272 ValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeu 291
Db      856 ATCTTGAAAGGTGTCCTTATATATGTTTACACCGCAAGTACTTGTGTTCCCGCTTG 915
Qy      292 IleAspLeuAlaIleAlaArgAsnThrLeuIleGlyLysAspAspPheLysSerLysGln 311
Db      916 GTTCAGCTGGCTGACACATGAGGATCATTCATTGGCGGAGACATCTCAAGTCGGAGCA 975
Qy      312 ThrLysMetLysSerValLeuValAspPheLeuValGlyAlaGlyLysProThrSer 331
Db      976 ACCAAGTTGAAGTCTGTTCTGGCCAGTTCTTATGATGATGACGATTAACCGGCTCTCC 1035
Qy      332 IleValSerTyrAsnHisLeuGlyLysAsnAspGlyMetAsnLeuSerAlaProGlnThr 351
Db      1036 ATTGCATCTTAAACCATTTAGGCAATATGACGCTTAACTTATCTGCTCCAAACAA 1095
Qy      352 PheArgSerLysGluIleSerLysSerAsnValValaAspMetValAsnSerAsnAla 371
Db      1096 TTATAGCTTAAGAGATTTCCAAAGTTCTGTGATAGATGACATCATCGCTTAATGAT 1155
Qy      372 IleLeuTyrGluPro-----GlyGluHisProAspHisValValIleLysTyrVal 389
Db      1156 ATCTTGTCATATGATTAACCTGGGTAAACAAAGTTGACCACTGATGTCTATCAATATATG 1215
Qy      390 ProTyrValGlyAspSerLysArgAlaMetAspGlyTyrThrSerGluIlePheMetGly 409
Db      1216 AAGCCCGTGGGAGCTCAAAAGTGCAAATGACGAGTATTCAGAGAGTGTATGTTAGT 1275
Qy      410 GlyLysSerThrIleValLeuHisAsnThrCysGluAspSerLeuLeuAlaAlaProIle 429
Db      1276 GGCCATTAACCGGATTTCAATTCACATGTTGGCAAGATTTCTTTACTGCTGACGCCCTTG 1335
Qy      430 IleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGluPheLys----- 446
Db      1336 ATCATCATCTTTTATGCTATGACGTAGTTTGTATCAAGATGCTCTATTAAGAGGTGAC 1395
Qy      447 -----AlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSer 464
Db      1396 CCAAGTTAAAGAAAGATGCGCAAAATTCAGAACTTTTATCCAGTTTAACTTCTTGAGT 1455
Qy      465 TyrLeuThrLysValaProLeuValProProGlyThrProValValAsnAlaLeuSerLys 484
Db      1456 TACTGTTAAAGCTCCATTAACAAAGACCGAGATTTCACCGGTTAATGCTTAAACAG 1515
```



```
QY 485 GlnArgAlaMetLeuGluIleMetArgAlaCysValGlyLeuAlaProGluIleAsn 504
Db 1519 CAGCGGACCGCTCTGGAAAACTTCTCGCTCGCTGATCGCGCTGCCGTCACGAACGAA 1578
QY 505 MetIleLeuGlu 508
Db 1579 CTGCGTTTCGAA 1590

RESULT 10
US-09-248-796A-3131
/ Sequence 3131, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Kelch Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 3131
/ LENGTH: 1578
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-3131

Alignment Scores:
Pred. No.: 7,31e-163 Length: 1578
Score: 1376.50 Matches: 263
Percent Similarity: 71.63% Conservative: 98
Best Local Similarity: 52.18% Mismatches: 128
Query Match: 52.30% Gaps: 15
Db: 4 Gaps: 5

US-10-718-952-2 (1-510) x US-09-248-796A-3131 (1-1578)
QY 17 GluThrGluIleGlnSerValTyrAsnTyrGluThrThGluLeuValIleGluAsnArg 36
Db 61 GAGCATCATTTATATACCAATTTACTTATGAAACTCT--GTGTGTTAAAAAGATGCT 117
QY 37 AaGgLTThrTyrGlnTrpIleValIysProIysSerValIysTyrGluPheIysThrAsn 56
Db 118 AATGCT-----AAATTCATCTGTTACACCAACGCTAGTACATGATTCATCAAGTGTAT 171
QY 57 ILeHsValProIysLeuGluIleValMetLeuValGlyTyrGlyIleYAsnAsnGlySerThr 76
Db 172 TTAAGAATCCCTTAAGGTGTTATATTGTTAGTCGATTCGGGTAAACAACGGTACTACA 231
QY 77 LeuThrGlyValIleAlaAsnArgGluGlyIleSerTyrAlaThrIysAspIysIle 96
Db 232 TTGTTAAGGCCACATTCACAGCAACAACAAATTTCTTTGAAAAACAAGAAGGTGTT 291
QY 97 GlnGlnAlaAsnTyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGly----- 114
Db 292 GTTAAACCAATTTACTATAGTCTGTCTGTACTCAAGCATTCACATCGTCAAAATCGGTGTGAT 351
QY 115 SerPheGlnGlyGluGluIleTyrAlaProPheIysSerLeuLeuProMetValAsnPro 134
Db 352 AAAGAAACCTGGTGAAGATGTTTACGTTTAACTTAACTTCATTTGTCATATGGTTAACCA 411
QY 135 AspAspIleValPheGlyGlyTyrPhePheIleSerAsnMetCAsnLeuAlaAspAlaMetAla 154
Db 412 AATGATTTGTGTTGATGTGGATATTAGTGTTCATTTACATTAACCAAGCAAGGAAG 471
QY 155 ArgAlaIysValPheAspIleAspLeuGluIysGlnLeuAspProTyrIleMetGluSerMet 174
Db 472 AAGAGCTAAAGTCTTGATGTACTTACAAACAGATGTATCATCACTATGGGAACAAG 531
QY 175 LeuProLeuProGlyIleTyrAspProAspPheIleAlaAlaAsnGlnGluIleArgAla 194
```

```
Db 532 AAACCTTTGGATCATCTACTACCTGATTTTCATTTGTAACCAAGTGAAGTCC 591
QY 195 AsnAsnValIleIleYs-----GlyThrIysGlnGluIleVal 206
Db 592 AACAAATGTTTTTAAACCAAGTCATATGTGAAGTTAAACTGATATATTAATGGCCACAGTT 651
QY 207 GlnGlnIleIleIysAspIleIysAlaPheIysGluAlaThrIysValAspIysValI 226
Db 652 GAAAAATACAGAAAAGATATCAGAGATTCAAGGCCAAACCAAGATTAAGATTAAGTATT 711
QY 227 ValLeuTyrThrAlaAsnThrGluArgTyrSerAsnLeuValIleGlyLeuAsnAspThr 246
Db 712 ATTTATATGACGCTTAATCTGAAGATATGCTATGTTTGCCAAATGTCAATACACT 771
QY 247 MetGluAsnLeuLeuAlaIleAspArgAsnGluIleGluIleSerProSerThrLeu 266
Db 772 GGTGCAACTTGATATCAAGATTTTAAAGATCACACGAAGAAATGCTCCAGTACTGTG 831
QY 267 TyrAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThr 286
Db 892 TTGCGGTTGCTTATCTGTGAAAAAGTTCCATATATTAATGTTTACACCAAAACACA 891
QY 287 PheValProGlyIleuIleAspLeuAlaIleAlaAsnThrLeuIleGlyIleYAspAsp 306
Db 892 TTGTTCCCGGTGTATTGATTAAGTCAAAAATACACTCATTCATTTGTTGTGATATAT 951
QY 307 PheIysSerGlyGlnThrIysMetIysSerValIleuValAspPheLeuValGlyIleGly 326
Db 952 TTCAAGTACAGCTCAACAAATAATTAATCAAGTTCAGTCAATTTGTTGTCAGTCTGCT 1011
QY 327 IleIysProThrSerIleValSerTyrAsnIleGluIysAsnAsnAspGlyMetAsnLeu 346
Db 1012 ATCAAAACCACTTATATGCTTCTTATATCACTTGCGGTAAACAATGACGTTACAAATTTA 1071
QY 347 SerAlaProGlnThrPheArgSerIysGluIleSerIysSerAsnValValAspAspMet 366
Db 1072 TCATCACCAAAAACAATTAAGATCTAAGAAAATTTCCAAACAATCTGTTGTGATATAT 1131
QY 367 ValAsnSerAsnAlaIleLeuTyr-----GluProGlyGluIleProAspIleValI 384
Db 1132 ATTGAAATCCAAAGAAATTTATTTGATACAAACAAGAAATGCTGACAAAGTTGATCTGATC 1191
QY 385 ValIleIysTyrValProTyrValGlyAspSerIysArgAlaMetAspGluTyrThrSer 404
Db 1192 GCATTTAATATCTGTCAGCTGTTGTTGATCTTAAGTTAGTCCATGATGAATACATATCA 1251
QY 405 GluIlePheMetGlyIleYsSerThrIleValIleuHsAsnThrCysGluAspSerLeu 424
Db 1252 GAATTAATGTTGGGTGCACAAACAATAAGTATTCACAAATGTTGGAAGATTCATTG 1311
QY 425 LeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGluLeuSerThrArgIleGlu 444
Db 1312 CTGCTACACCACTATATATATGACTTGTGTGCTGCTACTGAAATTTGCACTGAGTTCAA 1371
QY 445 PheIysAlaGluAsnGluIleYsPheIleSerPheIleSerPheIleProValAlaThrIleLeuSer 464
Db 1372 GTCAAAAGGCTCGGTAATCTGATATATGATGAATTTGATTCCTGTTGATGTTATTTGCT 1431
QY 465 TyrLeuThrIysValProLeuValProProGlyThrProValValAsnAlaLeuSerIys 484
Db 1432 TACTGCTCAAGCTCCATATAGCAAGACCAAGATTCAAACCATATCAACGATTAACAAA 1491
QY 485 GlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluIleAsn 504
Db 1492 CAAAGTCAACAAATTAAGTCAACTGCTTTCAGTGTGTTGTTTTCATTCATTAACCAACGA 1551
QY 505 MetIleLeuGlu 508
Db 1552 TTAAGATTTGAA 1563

RESULT 11
US-09-949-016-12608
```



```
Db 34522 GCGCACATTGAGTGGCGGAAGCCGGGTGAGCGGGTCAGGAGGGCCAGTCACAGA 34581
      : : : : :
Qy 260 ----- -GluIleSerProserT 265
Db 34582 CACAGCTGTGCTGACGTCGCCCTTGCCACAGCTCGGTGTGAGGGTGTGCCCTTCA 34641
      : : : : :
Qy 265 hLeuTyraIaIleAcyValMetGluAsnValProPheIleAsnGlySerProGlnA 285
      : : : : :
Db 34642 CGCTCTTGCGCGGTGCGACATCTTGAGGGGTGTGCTCTCATGTGGTCTCCGACAGA 34701
      : : : : :
Qy 285 enTh-PheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThreuIleGlyIYA 305
      : : : : :
Db 34702 ACACCTGTGTCGCCGAGCTCTTGAGCTCGGTGGCAGCACCGGGTTTTGTGGGCGGAG 34761
      : : : : :
Qy 305 sPAspPheIleSerGlyGlnThryMetIleSerValLeuValAspPheLeuValGlyYA 325
      : : : : :
Db 34762 ATGACTTCAGTCAAGTCAAGCCAGACCAAGTCAAGTCCGTGTGTGACTTCTCATTTGGCT 34821
      : : : : :
Qy 325 IaGlyIleIys----- 328
      : : : : :
Db 34822 CCGGGCTCAAGGTGGGTGGGCTGAGGAGCTGCCAGTGCAGAAAGGGGCTGGGGCGCG 34881
      : : : : :
Qy 328 ----- 328
      : : : : :
Db 34882 GAGCACTGTGCTGTGGGGCCGAGGGCTGTGAGCTGTGGGGCTTCCCTGTGACCCACA 34941
      : : : : :
Qy 329 --ProThSerIleValSerIYrAsnHlaLeuGlyAsnAsnAspGlyMetAenLeuSerA 348
      : : : : :
Db 34942 GACCATG-TCCATCGAGTTCACACACCTGGGCAACAACATGGGAGGAACCTTACGG 35000
      : : : : :
Qy 348 IaProGlnThrPheArgSerIYsGluIleSerIYsSerAsnValValAspAspMetValA 368
      : : : : :
Db 35001 CGCCATTGCAAGTTCGGCTCTAAGAGAGGTGTCCAAGAGCAACCGGTGAGCGACATGGTGC 35060
      : : : : :
Qy 368 snSerAsnAlaIleuTyrgIuProGlyGluHisProAspHisVal----- 383
      : : : : :
Db 35061 AGAGCAACCCAGTGTCTATACCGCGGAGAGAGCTGACACTG-CGTGCTGGGGCG 35119
      : : : : :
Qy 383 ----- 383
      : : : : :
Db 35120 CGGGCGCGGGCGGGGTTGCCCGGAGGGGTGGCAGGAGCCCGGCAACTCCTGCTCG 35179
      : : : : :
Qy 384 -----ValValIleIysIYrValProTYrValGlyAspSerIYsArgAlaMetAspG 401
      : : : : :
Db 35180 CACTCCAGGTGTCAATCAAGTATGTGCGTACGTGGGTCAACAGACCGCGGCTGGATG 35239
      : : : : :
Qy 401 IuTYrThSerGluIlePheMetGlyIYsSerThrIleValIleuHisAenThrCYsG 421
      : : : : :
Db 35240 ACTATACTCGGAGGTGATGTGGCGGGAACAACAACACTGGTGTGCAACACGTGTG 35299
      : : : : :
Qy 421 Iu----- 421
      : : : : :
Db 35300 AGGTGGCGGGCTTACAGGGCTGCTATGGAGGCGAGGCTTGGCCACCGCCCATCT 35359
      : : : : :
Qy 422 -----AspSerLeuLeuAlaIaProIleIleuAspLeuVal 434
      : : : : :
Db 35360 CCGAACCCACCGCCCGGAGGACTGTGCTGTGGCGGACCATCATGTGTGACCTAGCG 35419
      : : : : :
Qy 435 IeuleuAlaGluLeuSerThraIleGluPheIYsAlaGluAsnGluIYsPheHis 454
      : : : : :
Db 35420 CTGCTGACCGAGCTGTGCGAGGCGGTCTTGACATGCATGACATGCCACCGCGCAG 35479
      : : : : :
Qy 455 sArpPheIaProValAlaThrIleuSerIYrIeueThryAlaProIeuValProPro 474
      : : : : :
Db 35480 ACTTTCACCCCGGTGTGCTCGCTCACTTCCTTCAAGGCGCACTAGTGGCGGCC 35539
      : : : : :
Qy 475 GlyThrProValValAsnAlaLeuSerIYsGlnArgAlaMetLeuGluAsnIle----- 492
      : : : : :
Db 35540 GCGACCGCGGTGTCAATGCGCTTTCCGCCAGCGAGCTGCATCGAAGAACATCTCTCAGG 35599
      : : : : :
Qy 492 ----- 492
      : : : : :
```

```
Db 35600 TGCACCCAGGCTCCAGATTCCCGCCGAGGCCAGATATACCCCTGCTAAGCTATGCA 35659
      : : : : :
Qy 492 ----- 492
      : : : : :
Db 35660 CAGGCTTGGGGGCTGTAGGTGTGAGGAGATACCCCTTTCTCCGTGGGTGTCTGG 35719
      : : : : :
Qy 493 -----Met-ArgAlaCysValGlyLeuAlaProGluAsnMe 505
      : : : : :
Db 35720 CTCTCTAAGCTGTGTGACCTTCCAGGGGCTGTGGGGTCCCGCACAGAACACAT 35779
      : : : : :
Qy 505 rIleLeuGluTYrIYs 510
      : : : : :
Db 35780 GCTCTGGAACACAA 35795
      : : : : :

RESULT 12
US-09-397-787-34
; Sequence 34, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-34

Alignment Scores:
Pred. No.: 5.4e-128 Length: 1231
Score: 1098.50 Matches: 209
Percent Similarity: 81.29% Conservative: 56
Best Local Similarity: 64.11% Mismatches: 58
Query Match: 41.74% Indels: 3
DB: Gaps: 1

US-10-718-952-2 (1-510) x US-09-397-787-34 (1-1231)
Qy 187 AlaIaAsnGlnGluArgAla-AsnAsnValIleIYsGlyThryIYsGlnGluIYA 206
      : : : : :
Db 5 GCGGCAACCAAGCGCGCGCGGCAACCTCATCCAGGCTGCGTGGCGACAGCT 64
      : : : : :
Qy 206 IGlnGlnIleIleIYsAspIleIYsAlaPheIYsGluIaThryIYsValAspIYsValVa 226
      : : : : :
Db 65 GGAGCAGATCCGACGAGCATTCGAGACTTCCGTGTACCGCGGGCGTGGACAAAGTCA 124
      : : : : :
Qy 226 IValIeuTYrThraIaAsnThrGluArgTYrSerAsnLeuValValGlyLeuAsnAspTh 246
      : : : : :
Db 125 AGTGTGTGAGCGGGGACACGAGACGCTTGTGAGGTGATTCAGGCTTCACGACAC 184
      : : : : :
Qy 246 rMetGluAsnLeuLeuAlaIaValAspArgAsnGluIaGluIleSerProserThre 266
      : : : : :
Db 185 AGCCAGAACTGTGTGCGCACCATTTGAGCTCGGT--CTGAGGGTGTGCCCTTCACGCT 241
      : : : : :
Qy 266 uTYrAlaIleAcyValMetGluAsnValProPheIleAsnGlySerProGlnAsnTh 286
      : : : : :
Db 242 CTTCGCCGTGGCCAGCATCTTGAGGGGTGTGCTCTCATATGGGTCTCCGACAGAAC 301
      : : : : :
Qy 286 rPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThreuIleGlyIYsAsp 306
      : : : : :
Db 302 CCGTGTCCCGGAGCTTGTGAGCTGCGTGGCAGACACCGGGTTTTGTGGCGGAGATGA 361
      : : : : :
Qy 306 rPheIYsSerGlyGlnThryMetIYsSerValLeuValAspPheLeuValGlyAlaG 326
      : : : : :
Db 362 CTTCAAGTCAGGCGAGACCAAGTCMAGTCCGTGCTGTGTGACCTTCCATTGCTCCGG 421
      : : : : :
```



```
Db 19019 GAAGTACCAAGTCGCTGCTGACACCATCTG-----CAG 18981
Qy 376 ProGlyGluHisProAsp-----HisValValIleIleSerThrVal 389
Db 18980 CCCGACTGTACCCCGAGCTGTACAAGATACCGGCAAGAGTCCCATCTAC 18921
Qy 390 ProTyrValIleYAspSerThrArgAlaMetAspGluTyrThrSerGluIlePheMetGly 409
Db 18920 CCGCCCCGGGGGAGCCGGAAGAGGAGGTTGGGACACATGACATCACCGGATGGCTCGGC 18861
Qy 410 GlyLysSerThrIleValIleuHisAsenThrCysGluAspSerIleuValAlaProIle 429
Db 18860 TATCCGATCAGATCAAGTCAAGCTTCTCTGCGCGGACCTCATCTGCGCCCGCGCTG 18801
Qy 430 IleLeuAspLeuValIleuValIleuValIleuSerThrArgIleGluPheLysAlaGluAsn 449
Db 18800 GTCCTGTGACATGCGCTGTCTCTGACCTGCGCAGCGGCTGTGAGTGGCGGCGCATCCAG 18741
Qy 450 GluGlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyrIleuThrLysAla 469
Db 18740 GAGTGG-----ATGTCCTTCACTTCAAGACC 18714
Qy 470 ProLeuValProProGlyThrProValValAlaAlaLeuSerLysGluAlaAlaMetLeu 489
Db 18713 CCCATGCGCAGACCGGGCTCTGCGGCGGAGACAGACCTTCTCATCCAGCTCAACCAAGCTG 18654
Qy 490 GluAsnIleMetArgAlaCysValIleGlyLeuAlaProGluAsnAsnMetIleuGluTyr 509
Db 18654 AAGAACACCGCTGCGCTGCTGCGGGCGAGACCCCATCAACCACTCGGGCTCGACTAT 18594

RESULT 14
US-09-902-540-5256
; Sequence 5256, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Miesand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5256
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5256

Alignment Scores:
Pred. No.: 1,25e-57 Length: 1335
Score: 541.00 Matches: 140
Percent Similarity: 49.36% Conservative: 92
Best Local Similarity: 29.79% Mismatches: 190
Query Match: 20.55% Indels: 48
DB: 4 Gaps: 12

US-10-718-952-2 (1-510) x US-09-902-540-5256 (1-1335)
Qy 52 GluPheLysThrAsnIleHisValPro-----LysLeuGlyValMetLeuValGlyTyr 69
Db 4 GAAACAAAGAGGTCGGTCCGAAAGCCGAGCGCAAGCTCGCTCTGATTCGGGCTCG 63
Qy 70 GlyGlyAsnAsnGlySerThrLeuThrGlyGlyValIleAlaAsnArgLysGlyIleSer 89
Db 64 GCGCGCTGTCTCC--ACCAAGCTGATGCGCGGTGTGAGCTGCGCGGCAAGGCC----- 114
Qy 90 TrpAlaThrLysAspLysIleGlnGlnAlaAsnTyrPheGlySerLeuThrGlnAlaSer 109
```

```
Db 115 -----AAGGTACCCCATTTGGGTGCTACCGCATGGGC 150
Qy 110 AlaIleArgValGlySer---PheGlnIlyGluGluIleTyrAlaProPheLysSerLeu 128
Db 151 ACCGCCCGCTCTGGGGAAGCGAGCGGCGCACCGTG-----AACCTCAACAGAGCTG 204
Qy 129 LeuProMetValAsnProAspAspIleValPheGlyGlyTyrPAspIleSerAsnMetAsn 148
Db 205 GTGCTCTGGCGAGACTGAAGACGTTGCTTCCGCGCGCTGGAGCATATCCGTGAGAC 264
Qy 149 LeuAlaAspAlaMetAlaArgAlaLysValPheAspIleAspLeuGlnIleGlnLeuArg 168
Db 265 GCATACGAAGTGGCCGCTCGCGCGCTGCTACCGCAAGACCATGAGAGAGGTGAAG 324
Qy 169 ProTyrMetGluSerMetLeuProLeuProGlyIleTyrAspProAspPheIleAla 188
Db 325 CCGTTCCTCCAGAGATCAAGCGGAGAGGCGGTGCAAGACCTGATGTTGTCGCGCGC 384
Qy 189 AsnGlnGluGluArgAlaAsnAsnValIleLysGlyThrLys-----GlnGluGlnVal 206
Db 385 ATCGAGGCC-----AACCATCAAGGCCCAAGACCGCACCGCGAGAGCATC 432
Qy 207 GlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThrLysValAspLysValVal 226
Db 433 GAAGGCTGCGCCAGACATCCGCACTTCAAGAGAGCTTCAACCGACCGCGCCGCTG 492
Qy 227 ValLeuThrThrAlaAsnThrGluArgTyrSerAsnLeuValValIleGlyAsnAspThr 246
Db 493 ATGTGTGTGTGAGCAGCGGTGAGACCTTCGTCGCGCTCCCGATCTTCAAG---ACG 549
Qy 247 MetGluAsnLeuValAlaValAspArgAsnGluAlaGluIleSerProSerThrLeu 266
Db 550 CTGGCGCGCTTGAAGAGCGGCTGAGCAAGAACCCCGCATCAACCCACCGCGCTG 609
Qy 267 TyrAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySerProGln---Asn 285
Db 610 TACACCTTAGCGCCCATCAAGAGGCGGTGCTTCCGAGGACGCCACCGCCACCGCAC 669
Qy 286 ThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIleGlyLysAsp 305
Db 670 GTGACAGCGCGCGCTCCAGAGATGCGCAAGCAGAGATGCGTGTGCGCGCGCGCGC 729
Qy 306 AspPheLysSerGlyThrLysMetLysSerValLeuValAspPheLeuValAlaLysAla 325
Db 730 GACCTCAAGAGCGGCGCAGCATGTAAGACGTATCGCGCCGCTCAAGGCCCGCC 789
Qy 326 GlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsnAsnAspLysMetAsn 345
Db 790 ATGCTGGGTCTGATGCGTGTCTTCCACCAACATCTGTGGCAACCGGACCGGACAGTG 849
Qy 346 LeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSerAsnValValAspAsp 365
Db 850 CTGACAGACCGCGCGCTTCAAGGCCAAGAGATCAACCAAGTCAAGCTGTGACACCC 909
Qy 366 MetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisProAsp----- 381
Db 910 ATCTCG-----CAGCCCGACTGTACCCCGAGCTGTACAGAG 948
Qy 382 -----HisValValIleLysTyrValProTyrValIleGlyAspSerThrArgAlaMet 399
Db 949 TACGGCGCAAGGTGGCCCTTCACTATACCGCGCGGAGAGCGAGAGAGGTTGG 1008
Qy 400 AspGluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValIleuHisAsnThr 419
Db 1009 GACAACATCGACATACCGGAGTGGCTCGGTATCCGATGCGAGATCAAGTCACTTCTC 1068
Qy 420 CysGluAspSerLeuValAlaProIleIleLeuAspLeuValIleuValIleuVal 439
Db 1069 TCGCGGACTCATCTCGGCGCGCGGTGTGTGACATGCGCTGTCTGTGACCTG 1128
Qy 440 SerThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProVal 459
Db 1129 GCCAAGCGGCTGAGTGGCGGCGCATCCAGAGTGG----- 1164
```

QY 460 AlaThrIleuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValVal 479
DB 1165 -----ATGCTCTTACTTCAAGACCCCAAGCCGCGGCTCTGCGGTGAG 1215
QY 480 AsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIleuArgAlaCysValGlyLeu 499
DB 1216 CACGACTCTTCACTCAGCTCACCAAGCTGAGACACGCTGCGCTGCTGCGGCGAG 1275
QY 500 AlaProGluAsnMetIleuGlyTyr 509
DB 1276 GACCCCATCACCCCTCGGCTCGACTAT 1305

RESULT 15

US-09-313-294A-4684
; Sequence 4684, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4684
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 70034858H1
; NAME/KEY: unsure
; LOCATION: 185, 272, 275, 281, 283, 288
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4684

Alignment Scores:

Pred. No.:	4.23e-10	Length:	294
Score:	155.00	Matches:	27
Percent Similarity:	80.43%	Conservative:	10
Best Local Similarity:	58.70%	Mismatches:	9
Query Match:	5.89%	Indels:	0
DB:	4	Gaps:	0

US-10-718-952-2 (1-510) x US-09-313-294A-4684 (1-294)

QY 1 MetPheIleGluAsnPhelLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 125 ATGTCATCGAAGCTTCGCGTCAAGACCCCAAGCTGCGGCTCGGCGGAGATC 184
QY 21 GlnSerValTyrAsnTyrGlnThrThrGluLeuValIleGluAsnArgAsnGlyThrTyr 40
DB 185 NTCTCGAATACCGGTACGACACGACGAGCTGTGTACAGAGCCCAAGACGCGCTCC 244
QY 41 GlnTrpIleValLysPro 46
DB 245 CGCTGGTCTCGCGCCC 262

Search completed: June 7, 2005, 22:57:50
Job time : 310 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2005, 20:45:48 ; Search time 3675.5 Seconds

(without alignments)
860.161 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MFIEHFVKECPVXKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Egapop 6.0 , Egapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10718952/runat.06062005.173402.12940/app.query.fasta_1.1294
-DB=Published Applications NA -QMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=apct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USRR=US10718952.@CGN_1.1.879.@runat.06062005.173402.12940
-NCPU=6 -ICPU=3 -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2632	100.0	1533	14	US-10-025-003-9	Sequence 9, Appli
2	2632	100.0	1533	19	US-10-718-952-9	Sequence 9, Appli
3	2632	100.0	1760	14	US-10-025-003-1	Sequence 1, Appli
4	2632	100.0	1760	19	US-10-718-952-1	Sequence 1, Appli
5	2627	99.8	1533	14	US-10-025-003-5	Sequence 5, Appli
6	2627	99.8	1533	19	US-10-718-952-5	Sequence 5, Appli
7	2619	99.5	1533	14	US-10-424-599-70167	Sequence 70167, A
8	2600	98.8	1533	14	US-10-025-003-13	Sequence 13, Appli
9	2600	98.8	1533	14	US-10-025-003-15	Sequence 15, Appli
10	2600	98.8	1533	19	US-10-718-952-13	Sequence 13, Appli
11	2600	98.8	1533	19	US-10-718-952-15	Sequence 15, Appli
12	2593	98.5	1533	14	US-10-025-003-11	Sequence 11, Appli
13	2593	98.5	1533	19	US-10-718-952-11	Sequence 11, Appli
14	2524	95.9	2018	18	US-10-424-599-12022	Sequence 12022, A
15	2378	90.3	1533	9	US-09-938-842A-1438	Sequence 1438, Ap
16	2378	90.3	1533	11	US-09-938-842A-1438	Sequence 25858, A
17	2356	89.5	2156	18	US-10-425-114-25858	Sequence 47189, A
18	2356	89.5	2284	20	US-10-425-115-47189	Sequence 14714, A
19	2355	89.4	1924	19	US-10-767-701-14714	Sequence 2091, Ap
20	2354	89.3	2444	20	US-10-739-930-2091	Sequence 122151, A
21	2351	89.3	2444	20	US-10-425-115-122151	Sequence 70769, A
22	2347	89.2	2576	19	US-10-437-963-70769	Sequence 7081, Ap
23	2345	89.1	1898	18	US-10-425-114-7081	Sequence 10, Appli
24	2336	88.8	1931	9	US-09-921-232-10	Sequence 10, Appli
25	2336	88.8	1931	9	US-09-921-330-10	Sequence 5704, Ap
26	2336	88.8	1931	8	US-10-425-114-5704	Sequence 5612, Ap
27	2335	88.7	1879	18	US-10-425-114-5704	Sequence 1, Appli
28	2335	88.7	1914	18	US-10-425-114-5612	Sequence 26983, A
29	2335	88.7	1959	9	US-09-727-628-1	Sequence 47194, A
30	2335	88.7	1971	18	US-10-425-114-46983	Sequence 8483, Ap
31	2335	88.7	2233	20	US-10-425-115-47189	Sequence 47190, A
32	2335	88.7	2203	20	US-10-425-115-47189	Sequence 205, App
33	2334	88.7	1912	18	US-10-425-114-4483	Sequence 85285, A
34	2334	88.7	1997	20	US-10-425-115-47190	Sequence 16524, A
35	2283	86.7	1759	19	US-10-363-829-205	Sequence 5854, Ap
C 36	2205.5	83.8	2196	15	US-10-437-963-85285	Sequence 83553, A
37	2193	83.3	1853	18	US-10-425-114-16524	Sequence 47182, A
38	2193	83.3	2050	18	US-10-425-114-16009	Sequence 12021, A
39	2193	83.3	2193	18	US-10-425-114-5864	Sequence 12021, A
40	2193	83.3	2130	20	US-10-425-115-83553	Sequence 15, Appli
41	2129	80.9	1673	20	US-10-425-115-47180	Sequence 5, Appli
42	2104	79.9	1789	20	US-10-425-115-47182	Sequence 5, Appli
43	2023	76.9	2582	18	US-10-424-599-12021	Sequence 12021, A
44	1942.5	73.8	2582	18	US-10-424-599-12021	Sequence 12021, A
45	1813	68.9	3546	9	US-09-921-232-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997

PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-9

Alignment Scores:
Pred. No.: 1.15e-263 Length: 1533
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-718-952-2 (1-510) x US-10-025-003-9 (1-1533)

```
QY 1 MetPheIIeGluAsnPhelYValGluCysProAsnValIleTyrThrgIuThrgIuIle 20
DB 1 AAGTTCATCGAAGAATTAAAGTTGAGTCTCCATGTAAGTACACCGAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrgIuLeuValHieGluAsnArgSngIyThrTyr 40
DB 61 CAGTCGGTGTACACATACGAAACCACTTGTTCACGAGAACGSAATGGCACTTAT 120
QY 41 GlnTTPIIeValIleYpProIySerValIleYTyrgIuPhelYThrAsnIleHieValPro 60
DB 121 CAGTGGATGTGCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCT 180
QY 61 LysIleuGluValMetIleuValGluTTPGIYGIYAsnSngIYSerThrIeuThrgIyGIY 80
DB 181 AATTTAGGGGTATAGCTGTGGGTGGGTGGAAACAACGGCTCAACCTCACCGGTGCT 240
QY 81 ValIleValAsnArgGluGluYIleSerTTPAlaThrIleYAspLysIleGluGluIleAsn 100
DB 241 GTTATTGCTTAACCGAGAGGCAATTCAATGGCTACAAAGACAAAGATTCAACAGCCAT 300
QY 101 TyrPheGlySerIeuThrgIuIleSerAlaIleArgValGlySerPheGluGluGlu 120
DB 301 TACTTTGGCTCTCCCTCACCAAGCTCAGCTATCCGAGTGGGTCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheYSerIeuIeuProMetValAsnProAspIleValPheGly 140
DB 361 ATCTATAGCCCACTTCAAGAGCTGCTTCCAATGTGTTAACCTGACGACATGTGTGGG 420
QY 141 GlyTTPAspIleSerAsnMetAsnIleuValAspAlaMetAlaArgAlaIleYsValPheAsp 160
DB 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGCAAGGTGTTGAC 480
QY 161 IleAspLeuGluYsGluIleuArgProTyrMetGluSerMetIeuProIeuProGlyIle 180
DB 481 ATCGATTTTCAGAAAGCAGTGGAGGCTTACATGAAATTCATGCTCCCGGGAATC 540
QY 181 TyrAspProAspPheIleIleAlaAsnGluGluIleuArgAlaAsnAsnValIleYsGly 200
DB 541 TATGACCCGGATTTCTTGTCTGCAACCAAGAGACGGTGCACCAACGCTCAACAGGGGC 600
QY 201 ThrIySngIuGluIleuValGluIleIleIleYsAspIleYsValPheYsGluIleIle 220
DB 601 ACAAGCAAGAGAGAAATTCACAAATCATCAAGACATCAAGCCGTTTAAAGAGCAAC 660
QY 221 LysValAspLysValValIleuTTPThraIAsnThrgIuArgTyrSerAsnIleuVal 240
DB 661 AAAGTGAACAAGGCTGTACTGTGACTGCCAACACAGAGAGGATCAATATTGTTGTT 720
QY 241 ValGlyIleuAsnAspThrMetGluAsnIleuValIleValAspArgSngIuIleGlu 260
DB 721 GTGGGCTTTAAAGACCAATGAGAAATCTTGGCTGTGTGACGAATAGAGGTGAG 780
QY 261 IleSerProSerThrIeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
```

```
DB 761 ATTTCTCTCCATCCACTTGTATGCCATCTGTGTATGAAATGTTCTTCTATTAAT 840
QY 281 GlySerProGluAsnThrPheValProGlyIleuIleAspIleuAlaIleAlaArgSnrhr 300
DB 841 GGAAAGCCCTCAGAAACATTTGTACCAAGGCTGATGTGATGCTTGGCATGGCAAGAACT 900
QY 301 LeuIleGlyIleYsAspPheYsSerGlyIleThrIleYsMetIleYsSerValIleuValAsp 320
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTCTAGACCAAAATGAATCTGTGTGTTGAT 960
QY 321 PheIleuValGlyIleGlyIleYsProThrSerIleValIleSerTyrAsnHieIleuGlyAsn 340
DB 961 TTCTTGTGGGGGCTGTATGATCAAGCAACATCTATAGTCAAGTACCAACATCTGGGAAC 1020
QY 341 AsnAspGlyMetAsnIleuSerIleProIleThrPheArgSerIleGluIleSerIySer 360
DB 1021 AATGATGGTATGAACCTTTCGGCTCCACAACTTTCCTTCAGAGAAATCTCCAAAGAC 1080
QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleIleuTyrGluProGlyGluIleAspPro 380
DB 1081 AACGTTGTGATGATATGATGTCACACGCAATGCCATCTTATGAGCCCTGGTGAATCA 1140
QY 381 AspHieValIleValIleYsTyrValProTyrValGlyAspSerIleArgAlaMetAsp 400
DB 1141 GACCATGTTGTGTATTAAGATGTGCTTACGTAGGGGACAGCAAGAGCCATGAT 1200
QY 401 GluTyrThrSerGluIlePheMetGlyGlyIySerThrIleValIleuHieAsnThrCys 420
DB 1201 GAGTACACTTCAGAAATATCAATGGGTGAAAGACCACTTGTTCACAAACATGC 1260
QY 421 GluAspSerIleuIleAlaProIleIleuAspIleuValIleuIleuSer 440
DB 1261 GAGGATTCCTCTTACGCTGCTCTTATCTTGAGCTTGCTCTTCTTGAGCTGAGC 1320
QY 441 ThrArgIleGluPheIleYsGluAsnGluGlyIleYsPheHieSerPheHieProValIle 460
DB 1321 ACTAAGATCGAGTTTAAAGCTGAAATAGGGAATTCATCACTCCACCCAGTTGCT 1380
QY 461 ThrIleIleuSerIleuThrIleYsAlaProIleuValProProGlyThrProValIleAsn 480
DB 1381 ACCATCTTCAGTACTTCAACCAAGGCTCTCTGTGTTCAACGGGTACACCACTGGTGAAT 1440
QY 481 AlaIleuSerIleGluArgAlaMetIleuGluAsnIleMetArgAlaCysValGlyIleuAla 500
DB 1441 GCATTGTCAAGCAACGCTGCAATGCTGSAAAACATAAAGAGGCTGTGTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleuGluTyrIlys 510
DB 1501 CCAGAGAATTAACATGATTTCTCGAGTACAAAG 1530
```

RESULT 2
US-10-718-952-9
Sequence 9, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Strelitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
PRIORITY FILING DATE: 2003-11-21
PRIORITY FILING DATE: 08/835,751
PRIORITY FILING DATE: APRIL 8, 1997
PRIORITY FILING DATE: PCT/US98/06822
PRIORITY FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 1533

TYPE: DNA
ORGANISM: Glycine max
US-10-718-952-9

Alignment Scores:

Pred. No.:	1.15e-263	Length:	1533
Score:	2632.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
	19	Gaps:	0

US-10-718-952-2 (1-510) x US-10-718-952-9 (1-1533)

```

QY      1 MetPheIleGluAsnPhelysValGluCyProAsnValIysTyrThrGluThrGluIle 20
DB      1 ATGTCATCGAGAAATTTTAAGTTGAGTGTCTTAAGTGAAGTACCGGAGCTGAGATT 60
QY      21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB      61 CAGTCCGTTGACACTACGAAACACCAACGCTGTTCACGAGAACAGAAATGGCACTTAT 120
QY      41 GlnTyrIleValIysProIysSerValIysTyrGluPhelysThrAsnIleHisValPro 60
DB      121 CAGTGGATTGTCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCC 180
QY      61 LysLeuGlyValMetLeuValGlyTyrGlyIysAsnAsnGlySerThrLeuThrGlyGly 80
DB      181 AAATTGGGGTAAATGCTTGGGTGGGTGGGTAACACCGCTCAACCTCCACGGTGGT 240
QY      81 ValIleAlaAsnArgGluGlyIleSerTyrAlaThrIysAsnIleGlnIleAsn 100
DB      241 GTTATGTCATACCGAAGGGCATTTCATGGGCTACAAAGCAAGATTCACCAAGCCAA 300
QY      101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGlu 120
DB      301 TACTTTGGGCTCCCTCCACCAAGCTCAGCTATCCGAGTTGGGCTTCCACGGAGAGAA 360
QY      121 IleTyrAlaProPhelysSerLeuLeuProMetValAsnProAsnAspIleValPheGly 140
DB      361 ACTATAGCCCATTCAGAGCCCTGCTCCATGTTAAACCTGACGACATGTTGTTGGG 420
QY      141 GlyTyrAsnIleSerAsnMetAsnLeuAlaAsnAlaMetAlaArgAlaIysValPheAsp 160
DB      421 GGATGGGAATATCGACACATGAACCTGGCTGATGCGACAGGGCAAAAGGTTTGAC 480
QY      161 IleAspLeuGlnIysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
DB      481 ATCGATTTGCAGAACGAGTTGAGGCTTACATGGAATCCATGCTCCCGGAAATC 540
QY      181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleYsgly 200
DB      541 TATGACCCGGAATTCATGCTGCGCAACAGAGAGCGTCCAAACAGTCATCAAGGGG 600
QY      201 ThrIysGlnGluGlnValGlnGlnIleIleIysAspIleValAlaPheIysGluAlaThr 220
DB      601 ACAAAGCAAGAGAGCAAGTTCACCAATCATCAAGACATCAAGCGTTTAAAGAAAGCCAC 660
QY      221 LysValAspLysValValIleuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB      661 AAAGTGGACAAAGTGTGTGTACTGTGACCTGCCAACACAGAGAGGTAATTTGGTT 720
QY      241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
DB      721 GTGGGCTTAATGACACCATGAGAAATCTCTTGGCTGTGACAGAAATGAGCGTGAG 780
QY      261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB      781 ATTTCTCTTCCACCTTGATGCAATGCTTGTTGTAAGAAATGTCTTTCATTAAAT 840
QY      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB      841 GGAAGCCTCAGAACCTTTTGTACCAAGGCTGATTGATCTTCCATCCGAGAGAACT 900

```

```

QY      301 LeuIleGlyIysAspAspPheIysSerGlyGlnThrIysMetIysSerValLeuValAsp 320
DB      901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGACCAAAATGAAATCTGTGTTGAT 960
QY      321 PheLeuValGlyAlaGlyIleIysProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
DB      961 TTCCTTGTGGGGCTGTGATTCAGCCAAACATCTATATGTCAGTTACCAACATCTGGAAAC 1020
QY      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIysGluIleSerIysSer 360
DB      1021 AATGATGTATGAATCTTCCGCTCCACAACTTCCGTTCCAGAGAAATCTTCAAGAGC 1080
QY      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB      1081 AACGTTGTGATGATATGTGTCAACAGCAATGCAATCTCTATAGCCTGTGTAACTTCA 1140
QY      381 AspHisValValIleIysTyrValProTyrValGlyAspSerIysArgAlaMetAsp 400
DB      1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGCCATGAT 1200
QY      401 GlyTyrThrSerGluIlePheMetGlyGlyIysSerThrIleValIleuHisAsnThrCys 420
DB      1201 GAGTACACTTCAGAGATTCATGCGTGGAAAGACACCATTTGTCACAAACATGTC 1260
QY      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB      1261 GAGGATTCCTCTTACGCTCTCTTATATCTTGAGCTTGTCTTCTTGAGCTCAC 1320
QY      441 ThrArgIleGluPhelysAlaGluAsnGluGlyIysPheHisSerPheHisProValAla 460
DB      1321 ACTAGATGAGTTTAAGCTGAAAATGAGGAAAATTCATCATCTTCCACCACTGTTGCT 1380
QY      461 ThrIleLeuSerTyrLeuThrIysAlaProLeuValProProGlyThrProValValAsn 480
DB      1381 ACCATCTCAGCTACCTCACCAAGGCTCCTGTTTCCACCGGTACACCAAGTGGAAAT 1440
QY      481 AlaLeuSerIysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB      1441 GCATTGTCAACACAGGTGCAATGCTGGAACCATATATAGGGCTTGTGTGATTGGCC 1500
QY      501 ProGluAsnAsnMetIleLeuGluTyrIys 510
DB      1501 CCAAGAAATACATGATCTCCAGATACAG 1530

```

RESULT 3
US-10-025-003-1
; Sequence 1, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-1
Alignment Scores: 1.41e-263 Length: 1760

Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-718-952-2 (1-510) x US-10-025-003-1 (1-1760)

QY 1 MetPhe11eGluAsnPhelYsVal1GluCysProAsnVal1YsThrGluThrGlu1 20
DB 54 ATGTTTCATGAGAAATTTTAAGTGGTGTCTTAAGTGAATGACCCAGATGAGATT 113
QY 21 GlnSerVal1YsAsnTyrGluThrGluLeuVal1H1eGluAsnArgAsnGlyThrTyr 40
DB 114 CAGTCGCTGATACAACTACGAAACCAACGAACTTGTTCAGAGAAACGAAATGCACTTAT 173
QY 41 GlnTPr11eVal1YsProLysSerVal1YsTyrGluPheLysThrAsn11eH1sValPro 60
DB 174 CAGTGAATGTCAGAAACCAATCTGTCAAAATGCAATTTAAACCAACATCCATGTTTCT 233
QY 61 LysLeuG1YValMetLeuVal1G1YTrpG1Yg1YAsnAsnG1YSerThrLeuThrG1Yg1Y 80
DB 234 AAATTTAGGGGTAAATGTTTGGGTGGGTGGGTGAAACACGCTCAACCTTCAACGGGTGAT 293
QY 81 Val11eAlaAsnArgGluGly11eSerTPr1aThrLysAspLys11eGlnGlnAlaAsn 100
DB 294 GTTATGCTACCGAAGGCAATTTTCATGGCTACAAAGACAAAGATTCACCAAGCCAT 353
QY 101 TyrPheG1YSerLeuThrGlnAlaSerAla11eArgVal1G1YSerPheGlnG1Yg1Y 120
DB 354 TACTTTGGTCTCCCTCAAGGCTCAGTATCCGATTCGGTGGCTTTCAGAGAGAGAA 413
QY 121 11eTyrAlaProPheLysSerLeuLeuProMetVal1AsnProAspAsp11eVal1PheG1Y 140
DB 414 ATCTAAGCCCATTCAGAGCCCTGCTTCCAATGTTTAACTTCAGACATGTTGTTGG 473
QY 141 G1YTrpAsp11eSerAsnYsAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 474 GGATGGATATCGACAACTGAACCTGGCTGATGGCATGGCCAGGCAAAAGGTTTGAC 533
QY 161 11eAspLeuGlnLysGlnLeuArgProTyrMetGlnSerMetLeuProLeuProG1Y11e 180
DB 534 ATCGAATTCAGAGAGAGTTGAGGCTTTCATGGAATTCATGCTTCCACTCCCGGAATC 593
QY 181 TyrAspProAspPhe11eAla1aAsnGlnGluArgAlaAsnAsnVal11eLysG1Y 200
DB 594 TATGACCCGGATTTCTTGTCTGCCAACAGAGAGCGTCCACACATCTCATCAAGGAC 653
QY 201 ThrLysGlnGluGlnValGlnGln11eLysAsp11eLysAlaPheLysGluAlaThr 220
DB 654 ACAAGCAAGAGACAACTTCAACAAATCATCAAGACATCAAGCGCTTAAAGAGCAACC 713
QY 221 LysValAspLysVal1Val1LeuTPr1aThrAlaSerThrGluArgTyrSerAsnLeuVal 240
DB 714 AAAGTGAACAAGGTGTTTACTGCTGACTGCCAACAGAGAGCGTACGTAATTTGGT 773
QY 241 Val1G1YLeuAsnAspThrMetGlnAsnLeuLeuAla1ValaAspArgAsnGluAlaGlu 260
DB 774 GTGGGCTTAAATGACCACTGGAATCTTGGCTGCTGCTGAGCAAGAAATGAGGCTGAG 833
QY 261 11eSerProSerThrLeuTyrAla11eAlaCysValMetGluAsnValProPhe11eAsn 280
DB 834 ATTTGCTTCCACCTTGTATGCAATTTGCTGTGTTATGGAATAATCTTCTTCAATTAAT 893
QY 281 GlySerProGlnAsnThrPheValProG1YLeu11eAspLeuAla11eAlaArgAsnThr 300
DB 894 GGAAGCCCTCAGAAACCTTTTATACAGAGGCGATTTGATCTTCCATCCCGAGAACACT 953
QY 301 Leu11eG1Yg1YAspAspPheLysSerG1Yg1YThrLysMetLysSerVal1LeuValAsp 320
DB 954 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAGAACCAAAATGAAATCTGTGTTGGTGA 1013
QY 321 PheLeuValG1YAlaG1Y11eLysProThrSer11eValaSerTyrAsnH1sLeuG1YAsn 340

DB 1014 TTCTCTTGGGGGCGGTATCAAGCAACATCTAATAGTCAGTTACCAACATCTGGGAAC 1073
QY 341 AsnAspG1YMetAsnLeuSerAlaProGlnThrPheArgSerLysGlu11eSerLysSer 360
DB 1074 AATGATGATGAAATCTTTCGGCTTCAAAACTTTCGTTCCAGAAATCTCCCAAGAC 1133
QY 361 AsnVal1aAspAspMetValaAsnSerAsnAla11eLeuTyrGluProG1Yg1YH1sPro 380
DB 1134 AACGTTGTGATGATATGTCACAGCAATGCCATCTCTATGAGCCGTGTGAATCCA 1193
QY 381 AspH1sVal1Val11eLysTyrValProTyrVal1G1YAspSerLysArgAlaMetAsp 400
DB 1194 GACCATGTTGTTTATTAAGATGTGCTTACGTAAGGGAGACAGCAAGAGCCATGAT 1253
QY 401 GlnTyrThrSerGlu11ePheMetG1Yg1YLysSerThr11eValLeuH1sAsnThrCys 420
DB 1254 GAGTACACTTCAGAAATATTCATGGGTGAAGACACCACTTGTTCACACACATGC 1313
QY 421 GlnAspSerLeuLeuAla1aPro11eLeuAspLeuVal1LeuLeuAlaG1YLeuSer 440
DB 1314 GAGATTTCCCTTATGCTGCTCTATTAATCTTGAGCTTGGCTCTTGTGAGCTCAGC 1373
QY 441 ThrArg11eGluPheLysAlaGluAsnGluG1YLysPheH1sSerPheH1sProVal1a 460
DB 1374 ACTAAGATCGAGTTTAAAGCTGAATAATGAGGAAATTCACCTCATCCACCACTTCT 1433
QY 461 Thr11eLeuSerTyrLeuThrLysAlaProLeuValProProG1YThrProVal1aAsn 480
DB 1434 ACCATCTCCAGTACTTACCAAGGCTCTCTGTTCCACCGGATGACCACTGATGAT 1493
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsn11eMetArgAlaCysVal1G1YLeuAla 500
DB 1494 GCATTGTCAAGCAGCGTCAATGCTGGAATAACATTAAGAGGCTTGTGATTTGCC 1553
QY 501 ProGluAsnAsnMet11eLeuG1YtYrLys 510
DB 1554 CCAGAGAAATTAATGATTTCTGAGTACAAAG 1583

RESULT 4

US-10-718-952-1
; Sequence 1, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-1

Alignment Scores:

Pred. No.: 1,41e-263 Length: 1760
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-718-952-2 (1-510) x US-10-718-952-1 (1-1760)

QY 1 MetPheIleGIuAsnPhelYsValGIuCySProAsnValIlyeThrThrGIuThrGIuIle 20
DB 54 AGGTTATGAGAAATTTTAAAGTTGAGTGTCTTAAAGTGAATGACCCGAGCTGAGATT 113
QY 21 GlnSerValTYrAsnTYrGIuThrThrGIuLeuValHleGIuAsnArgAsnGIYThrTYr 40
DB 114 CAGTCGGTGTACACGAAACCAACGAGCTTTCACGAGAACAGAAATGGACACTAT 173
QY 41 GlnTPPleValIlyeProIySerValIlyeTYrGIuPhelYsThrAsnIleHleValPro 60
DB 174 CAGTGATGTCAAAACCAATCTGTCAATAGAAATTTAAACCAACATCCATGTTCCCT 223
QY 61 IlyeLeuGIValMetIleuValGIYTPGIYGIYAsnAsnGIYSerThrIleuThrGIYGIY 80
DB 234 AAATTAGGGGTAAATGCTTGGGTGGGTGGGTAACACGCTCAACCCCTCACCGGTGGT 293
QY 81 ValIleAlaAsnArgGIuGIYIleSerTPAlaThrIlyeAspIlyeIleGIuIleAlaAsn 100
DB 294 GTTATTTGTACCGAGAGGGCATTTCAATGGCTACAAAGACAAAGATTCAACAGCCAAAT 353
QY 101 TYrPheGIYSerIleuThrGIuAlaSerAlaIleArgValGIYSerPheGIuGIYGIuIle 120
DB 354 TACTTGGCTCCCTCACCCAAAGCTCAGCTATCCAGTTGGGTCTTCCAGGGAGAGAA 413
QY 121 IlyeTYrAlaProPhelYsSerIleuLeuProMetValAsnProAspAlIleValPheGIY 140
DB 414 ACTTAAGCCCCCATTCAGAGAGCTGCTTCAATGGTTAACCCGTGACCAACATGTGGTTGGG 473
QY 141 GIYTPAPAlIeSerAsnMetAsnIleuAlaAspAlaMetAlaArgAlaIlyeValPheAsp 160
DB 474 GGAATGGAGATACGACAACTGAACCTGGCTGATGCCATGGCCAGGGCAAAAGTGTGGAC 533
QY 161 IlyeAspLeuGIuIlyeGIuLeuArgProTYrMetGIYSerMetIleuProIyGIYIle 180
DB 534 ATCGATTTTGCAGAGAGAGTTGAGGCTTTCATGGAATTCATGTTCCATCCCGGAAATC 593
QY 181 TYrAspProAspPheIleAlaAlaAsnGIuGIuArgAlaAsnAsnValIleIyGIY 200
DB 594 TATGACCCCGAATTTCTGCTGCCAACAGAGAGAGCGCCAAACGTCAATCAAGGGGC 653
QY 201 ThrIlyeGIuGIuValGIuGIuIleIleIlyeAspAlIleValAlaPheIyGIuAlaThr 220
DB 654 ACAAGACGAGAGCAATTCACAAATCATCAAGACATCAAGCGTTTAAAGAGGACACC 713
QY 221 IlyeValAspIlyeValValIleuThrAlaAsnThrGIuArgTYrSerAsnIleuVal 240
DB 714 AAAGTGCACAGGTGGTGTATCTGTGACCTGCCAACACAGAGAGGTATATTGGATT 773
QY 241 ValGIYLeuAsnAspThrMetGIuAsnIleuAlaIleValAspArgAsnGIuAlaGIu 260
DB 774 GTGGGCTTAAATGACACATGAGATCTTGGCTGTGTGACAGAAATGAGCGCTGAG 833
QY 261 IlyeSerProSerThrIleuTYrAlaIleAlaCySValMetGIuAsnValProPheIleAsn 280
DB 834 ATTTCTCTTCCACCTGTGATGCAATGCTGTGTATGAAAATGTCTTTCATTAAT 893
QY 281 GIYSerProGIuAsnThrPheValProGIYLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 894 GGAAGCCCTCACAACACTTTTGTACAGAGGCTGATTGATCTTGCCATCCCGAGGAACACT 953
QY 301 LeuIleGIYGIYAspAspPheIySerGIYGIuThrIlyeMetIlyeSerValIleuValAsp 320
DB 954 TTGATTTGGTGAAGTCACTTCAAGAGTGTGACAGACCAAAATCAATCTGTGTGGTGGAT 1013
QY 321 PheLeuValGIYAlaGIYIlyeIlyeProThrSerIleValSerTYrAsnHleuGIYAsn 340
DB 1014 TTCTCTTGTGGGGCTGGTATCAAGCCAACTCTATATGATTAACACCACTTGGGAAAC 1073
QY 341 AsnAspGIYMetAsnIleuSerAlaProGIuThrPheArgSerIyGIuIleIySer 360
DB 1074 AATGATGTGATGATCTTCTGAGCTCCAAACACTTTCGTTCCAAAGAAATCTTCAAGAGAC 1133

QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTYrGIuProGIYGIuHlePro 380
DB 1134 AACGTGTGTGATGATGTGTCACACAAATGCAATCTTATGAGCTGTGTAACATCCA 1193
QY 381 AspHleValIleValIleIyTYrValProTYrValGIYAspSerIyAspAlaMetAsp 400
DB 1194 GACCATGTTGTGTATTTAAGATATGTGCTTCAATAGGGGACAGCAAGAGCCATGGAT 1253
QY 401 GIuTYrThrSerGIuIlePheMetGIYGIYIySerThrIleValIleuHleAsnThrCyS 420
DB 1254 GAGTACACTTCAGAGATATTCATGCGGTGAAAGACACCATGTTTGCACAAACATCG 1313
QY 421 GIuAspSerIleuAlaAlaProIleIleAspLeuValIleuAlaGIuIleuSer 440
DB 1314 GAGGATTCCTCTAGAGCTCTCTTATATCTTGAGCTTGTCTTCTGTCTGAGCTCACG 1373
QY 441 ThrArgIleGIuPhelYsAlaGIuAsnGIYGIYIyPheHleSerPheHleAspValAla 460
DB 1374 ACTAGATGTGATTTAAAGCTGAAATGAGGAAATTCACATTCACATCCAGTGTGCT 1433
QY 461 ThrIleuSerTYrIleuThrIlyeAlaProIleuValProProGIYThrProValIleAsn 480
DB 1434 ACCATCTCAGCTACCTCACCAAGGCTCTGTGTTCCACCGGTACACCAAGTGGAT 1493
QY 481 AlaIleuSerIyGIuArgAlaMetIleuGIuAsnIleMetArgAlaCySValGIYLeuAla 500
DB 1494 GCAATGTCAAGCAGGTGCAATGCTGGAACATATATGAGGCTTGTGTGATTTGGCC 1553
QY 501 ProGIuAsnMetIleuGIuTYrIlye 510
DB 1554 CCAAGAAATACATGATTTCTCGATGACAG 1583

RESULT 5

US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1

GENERAL INFORMATION:

; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-5

Alignment Scores:

Pred. No.: 3,8e-263 Length: 1533
Score: 2627.00 Matches: 509
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 14 Gaps: 0

US-10-718-952-2 (1-510) x US-10-025-003-5 (1-1533)

QY 1 MetPheIleGIuAsnPhelYsValGIuCySProAsnValIlyeThrThrGIuThrGIuIle 20
DB 1 AGGTTATGAGAAATTTTAAAGTTGAGTGTCTTAAAGTGAATGACCCGAGCTGAGATT 60

```

QY 21 GlnSerValTyrAsnTyrGluThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGGGTACAACTACGAAACCAACGAACTTGTTCAAGAACAGAAATGGCACTTAT 120
QY 41 GlnTTPDLeuValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATGTGTAACCAACCAAACTGTCAATACGAATTTMAAACCAAACTTCATGTTCT 180
QY 61 LysLeuGluValMetLeuValGlyTTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTAGGGGTAAAGCTTTGGGTTGGGTGGAGTGAACACAGCTTCACCTTCACCGGTGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTPAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTTAACCGAGAGGCAATTTTCATGGGCTACAAAGACAAAGATTTCACAAAGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTGGCTCCCTCCACCAAGCCTCAGCTATCCGAGTGGGTCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCTTCAAGACCTGCTTCCAAATGTTTAACTTCAGACACTTGTGTTGGG 420
QY 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATACGACAACTGAACCTGGCTGATGCCATGCCAGGCAAAAGGTGTGGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTGTCAGAAAGCAGTTGAGGCTTTCATGGAATTCATGCTTCCCTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TATGACCCGGAATTTCTGCTGCTGCCAACCAAGAGAGCGTCCCAACACTTCATCAAGGCG 600
QY 201 ThrLysGlnGluValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAAGCAGAGCAAGATTCAACAAATCATCAAAAGATCATCAAGCGTTTAAAGAAAGCACCC 660
QY 221 LysValAspLysValValIleuTTPThrAlaAsnThrGluArgTyrSerLeuLeuVal 240
DB 661 AAAGTGGACAAAGTGTGTCTACTGTGAGCTGCCAACAGAGAGGATCAAGTAAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTAAATGACCACTGAGAAATCTTGGCTGCTGTGACGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTGCTCTTCCACCTGTATGCAATGCTGTGTGTATGGAATAATGTTCTTTCATTAAAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAAGCCCTCAGAAACCTTTGTTCACAGGCGTGAATGATCTTGCACATCCGAGGAACT 900
QY 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGGAGATACCTTCAGAGTGTGACAGCAAAATGAATCTGTGTTGGTTGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTCCTTGGGGGCTGTATCAAGCCAACTATATGTCAGTTACAAACCATCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGAGGTATGATCTTTCGGCTCCACAAACTTTCCTCCAAAGAAATCTCCCAAGAC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTGATGATAGTCAACAGCAATGCCATCTCTATGAGCGTGTGAACTCA 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400

```

```

DB 1141 GACCATGTTGTTGTTATTAAAGTATGTGCTTACAGGAGGACAGCAATAGAGCCATGGAT 1200
QY 401 GlnTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAAATTTTCAATGAGGTGAGAAAGACCACTTGTGTTGCAACACATGC 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTATGCTGCTCTTATTAATCTTGAGCTTGTCCTTCTGCTGAGCTCAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAAGATCGAGTTTAAAGCTGAAATAGAGGAAATTCACACTATTCACCCAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATCTCAGCTTACCTACCAAGGCTCTCTGTTCCACCGGATACACAGTGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTGTCAAGCAGACGTCGCAATGCTGGAAAAACATPATGAGGCTTGTGTGATGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
DB 1501 CCAGAAATAACATGATTCTCGAGTACAAAG 1530

RESULT 6
US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streic, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-5

Alignment Scores:
Pred. No.: 3.8e-263 Length: 1533
Score: 2627.00 Matches: 509
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 19 Gaps: 0

US-10-718-952-2 (1-510) x US-10-718-952-5 (1-1533)
QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTCATCTCAGAAATTTTAAAGGTGAGTGTCTTAATGTGAATACACGAGCTAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGGGTACAACTACGAAACCAACGAACTTGTTCAAGAACAGAAATGGCACTTAT 120
QY 41 GlnTTPDLeuValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60

```

Db 121 CAGTGGATTGTCAAACCAATCTGTCAAATACGAATTTAAACCAACATCATGTTCTT 180
QY 61 LysLeuGlyValMetLeuValGlyTTPGlyValAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AATATTAGGGGTAAATGCTTGCGGTGGGGTGAACACGCTCAACCTCCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGlyGlyIleSerTPAlaThrLysAspLysIleGlnIleAlaAsn 100
Db 241 GTTATTGCTAACGAGAGGCAATTCATGGGCTCAAGACAAAGATTCAACAGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGln 120
Db 301 TACTTGGCTCCCTCACCAGGCTACGCTATCCGAGTTGGGCTTCCAGGGAGAGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTAAGCCCATTCAGAGCCCTGCTTCCAAATGGTTAACTTCAGCAATGTGGTTGGG 420
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GGAATGGATATACGACAACTGAACCTGGCTGATGCCATGGCCAGGCAAGGTGTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
Db 481 ATCGAATTCAGAGAGCAAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGlyLysArgAlaAsnAsnValIleLysGly 200
Db 541 TATGACCCCGAATTCATGCTGCTCCAAACAGAGAGACGCTCCAAACATGCTCAAGGCG 600
QY 201 ThrLysGlnGlnAlaGlnIleIleLysAspIleLysValPheLysGlyAlaIleAsp 220
Db 601 ACAAGACAAAGACAACTCAACAAATCAACAAAGATCAAGGCTTTAAGAGAGCACCC 660
QY 221 LysValAspLysValValIleuTTPThrAlaAsnThrGlyLysArgIleAsnLeuVal 240
Db 661 AAGTCGACAAAGGTGTTGATGCTGATGCTGCCAACACAGAGAGCTAATTTGGATT 720
QY 241 ValGlyLeuAsnAspThrMetGlyAsnLeuLeuAlaValAspArgAsnGlyAlaGly 260
Db 721 GTGGGCTTAAATGACACCAATGAGAAATCTTGGCTGCTGAGCAAGAAATGAGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGlyAsnValProPheIleAsn 280
Db 781 AATTCCTCTTCCACTTGTATGCAATTCCTGTGTATGGAATATCTCTTCATTAAAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAGCCCTCAGAACCTTTGTACAGAGGCTGATTTGATCTTCCATCCGAGAGAACAT 900
QY 301 LeuIleGlyLysAspAspPheLysSerGlyInThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTGGTGGAGATACATTCAGAGTGGTCAGAACCAAAATCTGTGTGGTGGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleLeuGlyAsn 340
Db 961 TTCCTTGGGGGCTGGATCAAGCCAAATCTATAGTACATTCATCAACATCGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysSer 360
Db 1021 AATGATGCTATGATTTTCCGCTCCAAACTTTCCTCCAGAGAAATCTCCAGAGAC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlyLeuProGlyGlnHisPro 380
Db 1081 AACGTTGTGATGATGTGTCAAACAGCAATGCATCTCTAATAGCTGCTGTAACATCA 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCAATGTTGTATTAAGATATGCTTACGTAAGGAGACAAATAGAGCCATGAGAT 1200
QY 401 GlyTyrThrSerGlyLysPheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAGACCAATGTTTTCACAAACATATC 1260

QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGlyLeuSer 440
Db 1261 GAGGATTCCTCTTACTTACTCTCTATTAATCTTGAGACTTGTCTTCTGAGCTCAGC 1320
QY 441 ThrArgIleGlyPheLysAlaGlnAsnGlnGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAGAAATGAGTTTAACCTGAAAATGAGGAAAATTCACATCATTCCACCAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
Db 1381 ACCATCTCAGCTACCTCACCAAGGCTCTCTGTTCCACCGGGTACACAGTGGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GAATGTCAAGCAGCGTGCATATGCTGGAACCATTAATAGAGGCTTGTGTGATTTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
Db 1501 CCAAGAAATTAACATGATTTCTCAAGTACAG 1530

RESULT 7
US-10-424-599-70167
; Sequence 70167, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 70167
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1
US-10-424-599-70167

Alignment Scores:
Pred. No.: 3,81e-262 Length: 1989
Score: 2619.00 Matches: 510
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 99.51% Indels: 1
DB: 18 Gaps: 0

US-10-718-952-2 (1-510) x US-10-424-599-70167 (1-1989)

QY 1 MetPheIleGlyAsnPheLysValGlyCysProAsnValLysTyrThrGlnThrGlyIle 20
Db 217 ATGTTCAATCGAAGAAATTTAAAGTTAGGTGCTCTTAATGTGAAGATACCAAGACTGAGAT 276
QY 21 GlnSerValTyrAsnTyrGlyThrThrGlnLeuValHisGlnAspArgAsnGlyTyrTyr 40
Db 277 CAGTCCGTTACAACTACGAACCAACCGAACTTGTTCACGAGAAACGAAATGGCACTTAT 336
QY 41 GlnTPIleValLysProLysSerValLysTyrGlyLysPheLysThrAsnIleHisValPro 60
Db 337 CAGTGGATTGTCAAACCAATCTGTCAAATACGAATTTAAACCAACATTCATGTTCTT 396
QY 61 LysLeuGlyValMetLeuValGlyTTPGlyValAsnAsnGlySerThrLeuThrGlyGly 80
Db 397 AATATTAGGGGTAAATGCTTGCGGTGGGGTGAACACGCTTACCTTACCCGCTGAG 456
QY 80 ValIleAlaAsnArgGlyGlyIleSerTPAlaThrLysAspLysIleGlnIleAlaAsn 100
Db 457 TGTATTGCTAACAGAGAGGCAATTCATGGGCTACAAAGACAAAGATTCAACAAAGCCAA 516

QY 100 nTyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 517 TTACTTGGCTCCCTCACCACCAAGCCTCAGCTATCCAGTTGGGCTCTTCACAGGAGAGA 576
QY 120 uIleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheG 140
Db 577 AATCATGCCCCATTCMAAGCCTGCTTCCAAATGGTTAAACCTTGACGACATGTTGTTGG 636
QY 140 yGlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAs 160
Db 637 GGGATGGGATATCAGACATGAACTGGCTGATGCCATGGCCAGGCGAAAGGTGTTTGA 696
QY 160 pIleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyT 180
Db 697 CATCGATTTGGAGAAGCAGTTGAGGCTTACATGGAATCCATGCTTCCACTCCCGGAAAT 756
QY 180 eTyrAspProAspPheIleAlaIleAsnGlnGluValGlnAlaAsnValIleLysG 200
Db 757 CTATGACCCGGAATTTCAATGCTGCCAACAGAGAGCGCTGCCAACACGTCATCMAAGG 816
QY 200 yThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysValPheLysGlnAlaTh 220
Db 817 CACAAAGCAAGACAAAGTTCAACAAATCAACAAAGCATCAAGGCGTTTAAAGAAAGCCAC 876
QY 220 rLysValAspLysValValIleLeuTTPAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 877 CAAAGTGGACAAAGGTGGTTGTACTGTGAGCTGCCAACACAGAGAGGTACAGTATTTGGT 936
QY 240 lValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaAlaValAspArgAsnGluAgl 260
Db 937 TGTGGGCTTTAATGACACCATGGAGAAATCTTGGCTGTGGACAGAAATGAGGCTTA 996
QY 260 uIleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAs 280
Db 997 GATTTCTCCTTCCACTTGTATGCCATGTGCTGTGTAAGAAATGTTCTTCTTAA 1056
QY 280 nGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnTh 300
Db 1057 TGGAAAGCCTCAGAACACTTTTGTACAGAGGCTGATTGATCTTGGCATGGCGAGAACAC 1116
QY 300 rLeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAs 320
Db 1117 TTTGATTTGGTGGAGATGACTTCAAGAGTGTGTCAGACAAATGMAATCTGTGTGTTTA 1176
QY 320 pPheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHiIleGluGlyAs 340
Db 1177 TTTCCTTGGGGGCTGGTATCAAGCCAACTATATGATCAAGTTACAAACCATCTGGGAAA 1236
QY 340 nAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSe 360
Db 1237 CATATGATGGATTAATCTTTGGGCTCCACAACTTCCGTTCCAAAGAAATCTCCAAAG 1296
QY 360 rAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHiIsp 380
Db 1297 CAACGTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1356
QY 380 oAspHiIleValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAs 400
Db 1357 AGACCATGTTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1416
QY 400 pGluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHiIleAsnThrCy 420
Db 1417 TGAAGTCACTTCAGATATATTCATGGGTGGAAAGACACCATGTTTTCACAAACATAG 1476
QY 420 sGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSe 440
Db 1477 CGAGGATTCCTCTCTAGTGTGCTCTATATCTTGGACTTGGTCTCTTGGTGAAGCTCAG 1536
QY 440 rThrArgIleGluPheLysAlaGluLeuGlnGlyLysPheHiIleProValAl 460
Db 1537 CACTAAGATCGAATTTAAAGCTGAAAAATAGGAAAAATTCACACTATTCACACCAAGTTGC 1596

QY 460 aThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAs 480
Db 1597 TACCATCTCAGCTACCTCACCACAGGCTCTCTGTTCACACGGGTACACCAAGTGTGA 1656
QY 480 nAlaLeuSerLysGlnAlaGlnAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAl 500
Db 1657 TGCATTTGTCMAAGCAGCGTGCATCTGGAAACATTAATGAGGCTTGTGTGGATTGGC 1716
QY 500 aProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1717 CCCAGAGATATAATGATTTCTGAGTACAG 1747

RESULT 8
US-10-025-003-13
; Sequence 13, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Strelt, Leon
; APPLICANT: Streil, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-025-003-13

Alignment Scores:
Pred. No.: 2,476-260 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
Gaps: 0
DB: 14

US-10-718-952-2 (1-510) x US-10-025-003-13 (1-1533)

QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
Db 1 ATGTTCAATCGAAGAAATTTTAAGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGAT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 61 CAGTCGGGTACAACTACCAACCAACCACTTGTTCAGAGAAACGGAATGACACTAT 120
QY 41 GlnTPIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db 121 CAGTGGATTTGCAAAACCAATCCGTCAACTACATTTAAACCAACCAACCATGTTCA 180
QY 61 LysLeuGlyValMetLeuValGlyTTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AAATTGGGGGTATGCTTGTGGGTGGGAAACCAACGCGCTTACCTCAACCGGTGCT 240
QY 81 ValIleAlaAsnArgGluGlyLysSerTPAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 241 GTTATTTGCAACAGAGAGGCAATTCATGGGCTTACAAAGGAAAGCAATTCACAAAGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTTTGGGCTCCCTTACCCCAAGCCTCAGCTATTCAGATTTGATCTTCCAGGAGAGAGA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140


```

Db      361 ATCTATGCCCATCTCAAGAGCTGCTTCCAAATGTTATCTCAACACATGCTTGGG 420
Qy      141 GYTPAPSPILeserAenMeAenLeuAlaSPAlaMeAlaArgAlaLysValPheASP 160
Db      421 GGATGGGATATACGAACAATGAACTGGCTGATGCCATGGCCAGGGCAAGGTTGTTGAC 480
Qy      161 ILeaSPLeuGlnLysGlnLeuAArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db      481 ATCGATTTCAGAGAGAGTTGAGGCGCTTACATGAAATCCATGGTTCCATCCCCGGAATC 540
Qy      181 TYASPProASPheIleAlaAlaAenGlnGluLysArgAlaAenValIleLysGly 200
Db      541 TACGACCCGAGTTTCTTGTGCTGCCAACAGAGAGAGCTGCCAACAGCGATTAAAGGC 600
Qy      201 ThrLysGlnGlnGlnValGlnGlnIleIleLysASPLeuLysAlaPheLysGlnAlaThr 220
Db      601 ACAAGACCAAGACAAAGTTCAGCAAAATCATCAAGACATCAAGCGTTTAAAGAACCCACC 660
Qy      221 LysValASPLeuLysValValIleuTyrThrAlaAenThrGluArgTyrSerAenLeuVal 240
Db      661 AAAGTGAACAAGGTGTGTGCTGTGACTGCGCAACAAGAGAGGTATAGCAATTTGGTT 720
Qy      241 ValGlyLeuAaSPThrMetGluAenLeuAlaAlaValASPArgAenGlnAlaGlu 260
Db      721 GTAGGCTTATATACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
Qy      261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAenValProPheIleAen 280
Db      781 ATTTCTCTTCCACCTTGATATGCAATGCTGTGTATGGAATAATGTTCTTTCATTAAT 840
Qy      281 GlySerProGlnAenThrPheValProGlyLeuIleASPLeuAlaIleAlaAenThr 300
Db      841 GGAAGCCCTCAGAAACCTTTTGAACAAGGCTGATGATCTTCCATCGGAGAAACACT 900
Qy      301 LeuIleGlyLysASPheLysSerGlyGlnThrLysMetLysSerValIleuValASP 320
Db      901 TTGATTGGTGGAGATATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTTGAT 960
Qy      321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAenIleLysGlyAen 340
Db      961 TTTCTTGTGGGGCTGATCAAGCCCAACATCATATGTTAGTTACAAACCATCGGGAAC 1020
Qy      341 AenASPLeuMetAenLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db      1021 AATGAGGTATGATATCTCGGCTCCACAAACCTTCGCTCCAAAGAAATCTTCCAAAGAGC 1080
Qy      361 AenValValASPheMetValAenSerAenAlaIleLeuTyrGluProGlyGlnIleASP 380
Db      1081 AACGTTGTTGACATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGAAACATCCC 1140
Qy      381 ASPheIleValValIleLysTyrValProTyrValGlyASPheLysArgAlaMetASP 400
Db      1141 GACCATGTTGTGTATTAAGTATGCTTACGTAGGGGATGACAAAGAGAGCATGGAT 1200
Qy      401 GlyTyrThrSerGlnIlePheMetGlyLysLysSerThrIleValIleuIleAenThrCys 420
Db      1201 GAGTACACTTCAGAGATATTCATGGGTGAGAAAGAACCAATGTTTTCACCAACATAT 1260
Qy      421 GluASPheLysLeuAlaAlaProIleIleLeuASPheValIleuLeuAlaGlnLeuSer 440
Db      1261 GAGAGATCCCTTTTACCTGCTCTATTAATCTTGGACTGCTGCTTCTTGTGAGCTGAGC 1320
Qy      441 ThrArgIleGluPheLysAlaGlnAenGlnGlyLysPheIleSerPheIleProValAla 460
Db      1321 ACTAGAAATCCAGTTTAAAGCTGAATTAAGAGGAAATTCACATCTTCCACCCAGTTGGC 1380
Qy      461 ThrIleLeuSerTyrIleThrLysAlaProLeuValProProGlyThrProValValAen 480
Db      1381 ACCATTCTAGCTATATGACCAAGGCTCTGTGTTCCACCGGTACACACAGTGGGAAT 1440
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGluAenIleMetArgAlaCysValGlyLeuAla 500

```

```

Db      1441 GCATTGTCAAAGAGCGTGCATGCTGGAAGAAACATATATGAGGCTTGTGTGATTGGCC 1500
Qy      501 ProGluAenAenMetIleLeuGlnTyrLys 510
Db      1501 CCAGAGAAATATACATGATCTTCGAGTACAAG 1530

RESULT 9
US-10-025-003-15
/ Sequence 15, Application US/10025003
/ Publication No. US20030074685A1
/ GENERAL INFORMATION:
/ APPLICANT: Hitz, William
/ APPLICANT: Sebastian, Scott
/ APPLICANT: Grace, John
/ APPLICANT: Sreelit, Leon
/ TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
/ FILE REFERENCE: BB-1077-C
/ CURRENT APPLICATION NUMBER: US/10/025,003
/ PRIOR FILING DATE: 2002-05-07
/ PRIOR APPLICATION NUMBER: 08/835,751
/ PRIOR FILING DATE: APRIL 8, 1997
/ PRIOR APPLICATION NUMBER: PCT/US98/06822
/ PRIOR FILING DATE: APRIL 7, 1998
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 15
/ LENGTH: 1533
/ TYPE: DNA
/ ORGANISM: Glycine max
US-10-025-003-15

Alignment Scores:
Pred. No.: 2,47e-260 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
Gaps: 0
DB: 14

US-10-718-952-2 (1-510) x US-10-025-003-15 (1-1533)

Qy      1 MetPheIleGluAenPheLysValGlyCysProAenValLysTyrThrGlnThrGlnIle 20
Db      1 ATGTTATCTGAGAAATTTTAAGGTAGAGACTCTTAATGTAAGTACACCGAGACTGAGATT 60
Qy      21 GlnSerValTyrAsnTyrGlnThrThrGluLeuValIleGluAaSPArgAenGlyThrTyr 40
Db      61 CAGTCCGTTACAACTACGAAACCAACGAACTTTGTTACGAGAAACGAAATGGCACTTAT 120
Qy      41 GlnTPRIleValLysProLysSerValLysTyrGlnPheLysThrAenIleHISValPro 60
Db      121 CAGTGGATGTCAAACCCAAATCCGTCACACTACCAATTTAAACCAACCAACCATGTTCCA 180
Qy      61 LysLeuGlyValMetLeuValGlyTyrGlyAaenGlySerThrLeuThrGlyGly 80
Db      181 AAATTGGGGGTATGCTGTGGGTGGGATGAAACAGGCTCTTACCTTCAACCGGTGTGT 240
Qy      81 ValIleAlaAenArgGlnGlyLysSerTPRIleThrLysASPheIleGlnGlnAlaAen 100
Db      241 GTTATTGCTTACAGAGAGGCAATTCATAGGCTTACAAAGACAAAGATTCAACAGCAAT 300
Qy      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGln 120
Db      301 TACTTGGCTCCCTCACCACCAAGCTCAGCTATTCAGATGATGATCTTCCAGGAGAGGAA 360
Qy      121 IleTyrAlaProPheLysSerLeuLeuProMetValAenProASPheIleValPheGly 140
Db      361 ATCTATGCCCATCTCAAGAGCTGCTTCCAAATGTTTAAATCTTCCAGAGCAATTTGTTGGG 420
Qy      141 GYTPAPSPILeserAenMeAenLeuAlaSPAlaMeAlaArgAlaLysValPheASP 160
Db      421 GGATGGGATATACGAACAATGAACTGGCTGATGCCATGGCCAGGGCAAGGTTGTTGAC 480

```


161 ILeaSPLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
481 ATCGATTTCAGAAAGCAGTTGAGGCTTTCATGGAATTCATGTTCCATCCCGGAAATC 540
181 TyrAspProAspPheIleAlaIaAsnGlnGluArgAlaAsnAsnValIleIysGly 200
541 TACGACCCCGATTTCATGCTCTGCAACCAAGAGAGCGGCCAACACCTGATTGAAGGCG 600
201 ThrLysGlnGluValGlnGlnIleIleLysAspIleLysAlaPheLysGlnValThr 220
601 ACAAAGCAAGACAGATTGAGCAAAATCATCAAGACATCAAGCCGTTTAAGGAAGCAC 660
221 LysValAspLysValValIleLeuThrAlaAsnThrGluArgLysSerAsnLeuVal 240
661 AAAGGAGCAAGGTGGTGTCTCTGTCGACTGCAACACAGAGAGGATATGCAATTTGGTT 720
241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluValGlu 260
721 GTAGGCTTAATGACACCATGAGAAATCTCTGGCTGCTGTCGACAGAAATGAGGCTGAG 780
261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
781 ATTTCCTCTCCACCTTGATGCAATGCTGTGATGGAATAATGTTCTTTCATTAAAT 840
281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
841 GGAAGCCCTCAGAAACCTTTTGACAGAGGCTGATTGATCTTGCCATCCGAGGAACACT 900
301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
901 TTGATTGGTGGAGATGACCTTCAAGAGTGGTCAGAACCAAAATGAAATCTGTGGTGGAT 960
321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
961 TTTCCTGTGGGAGCTGGATCAAGCCAAATCATATGTTAGTTACAAACATCTGGGAAAC 1020
341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
1021 AATGATGGATGATATCTCGGCTCCAAACCTTCGCTCCAGGAATCTCCAAAGAAC 1080
361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyLysIlePro 380
1081 AACGTTGTGAGATATGCTCAACAGCAATGCCATCTCTATGAGCCTGTGAAACATCCC 1140
381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
1141 GACCATGTTGTTATTAAGTATGCTTACGTACGAGGATGACAAAGAGCATGAGAT 1200
401 GlyTyrThrSerGlnIlePheMetGlyLysSerThrIleValLeuHisAsnThrCys 420
1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAACACCATGTTTGGCAACACATGT 1260
421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
1261 GAGGATTCCTCTTAGCTCTCTATATCTTGAATCTGTCCTTCTCTGAGCTGAGAC 1320
441 ThrArgIleGluPheLysAlaGluAsnGlnGlyLysPheHisSerPheHisProValAla 460
1321 ACTAGATCAAGTTTAAGCTGAAATGAGGGAATTCACACTTCATCCACCCAGTTGCT 1380
461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
1381 ACCATTCTCAGCATATGACAAAGCTCTCTGTTCCACCGGATCACACAGTGGAGAT 1440
481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
1441 GGATTGTCAAAGCAGCGTCAATGCTGAAACATATAGAGGCTGTGTGATTGGCC 1500
501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
1501 CCAGAAATTAACATGATTCTCGAGTACAAG 1530

RESULT 10
US-10-718-952-13
; Sequence 13, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-13
Alignment Scores:
Pred. No.: 2,476-260 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
Gaps: 0
DB: 19
US-10-718-952-2 (1-510) x US-10-718-952-13 (1-1533)
QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTTCAATGCAAGAAATTTTAAGGATAGAGATCCTATATGGAATGATACCGAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCCGCTGATCACTACGAAACCAACCACTTGTTCAGAAAGAGATGGCACTTAT 120
QY 41 GlnTPHLeuValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATTGTCAAAACCAATCCGTCACATACCAATTTAAACCAACCACTTTCCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyLysAsnAsnGlySerThrLeuThrGlyIle 80
DB 181 AAATTGGGGGTGATGCTTGGGGTGGGATGGAACCAACGCTCTACCTCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGlnGlyLysSerThrAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATGCTCAAGAGAGGCACTTTCATAGGCTTACAAAGAGAAATTCACACAGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
DB 301 TACTTTGGCTCCCTACCCCAAGCCCTCAGCTATTCAGATTGGATCTTCCAGGAGAGAGA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTCAGAAAGTCTGCTTCCAAATGTTAATCTCGACGACATTTGTTGGG 420
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGATATCAAGAAATGAACTGCTGATGCAATGGCCAGAGGCAAGGTTTGAC 480
QY 161 ILeaSPLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTCAGAAAGCAGTTGAGGCTTTCATGGAATTCATGTTCCATCCCGGAAATC 540
QY 181 TyrAspProAspPheIleAlaIaAsnGlnGluArgAlaAsnAsnValIleIysGly 200

```

Db      541 TAGACCCGGATTCTTCTGCTGCAACAGAGACGCTGCCAACACAGCGATTAAAGGC 600
Qy      201 ThrTysGlnGlnValGlnGlnIleIleTyrAspIleValAlaPheTyrGlnValThr 220
Db      601 ACAAAGCAAGACAGATTGAGCAAAATCAACAAGACATCAAGCGTTTAAAGAAAGCACCC 660
Qy      221 LysValAspLysValValValLeuTrrPrrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      661 AAGAGTGAACAAGGTGGTGTCTCTGTGACTGCAACACAGAGAGGTATAGCAATTTGGT 720
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAlaAspArgAsnGluAlaGlu 260
Db      721 GTAGGCTTAAATGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGCTGAG 780
Qy      261 IleserProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      781 ATTTCTCTTCCACCTTGTATGCAATTCGCTGTGATGCAAAAATGTTCTTTCATTAAAT 840
Qy      281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      841 GGAAGCCCTCAGAAACCTTTGTACAGAGGCTGATTGATCTTGCCATCCGAGAGAACT 900
Qy      301 LeuIleGlyGlyAspAspPheTyrSerGlyGlnThrTyrMetTyrSerValLeuValAsp 320
Db      901 TTGATTGGTGGAGATGACTTCAAGACTGCTCAGACCAAAATGAAATCTGTGTGGTGGAT 960
Qy      321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleGluGlyAsn 340
Db      961 TTTCTTGTGGGGCTGTGTCAAGCCACATCATATGATTAGTTACAACCATCTGGGAAAC 1020
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGluIleSerLysSer 360
Db      1021 AATGATGATGATGATATCTCGCTCCCAACCTTCGCTCCCAAGAAATCTCCAGAGAC 1080
Qy      361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db      1081 AACGTTTGTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAACATCCC 1140
Qy      381 AspHisValValIleLysTyrValProTyrValGlyAspSerIleAspValAlaMetAsp 400
Db      1141 GACCATGTTGTGTATTAAGTATGTCCTTACGTAGGGGATGACAAAGAGAGCATGGAT 1200
Qy      401 GlnTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db      1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCAATGTTTTCACAAACATGT 1260
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1261 GAGGATTCCTTTTACCTCTCTCTATTAATCTTGAGCTTGCTCTTCTGCTGAGCTGAGC 1320
Qy      441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      1321 ACTAGATTCAGATTAACTGAAAATGAGGAAAATTCACACTCATTTCCACCACTGGT 1380
Qy      461 ThrIleLeuSerTyrLeuThrTyrAlaProLeuValProProGlyThrProValValAsn 480
Db      1381 ACCATTCTCAAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGTACACACAGGTGGAAAT 1440
Qy      481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1441 GCATTGTCAAGACAGGTCATGCTGAGAAAACATATATAGGGCTGTGTGTGGATTGGCC 1500
Qy      501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db      1501 CCAGAGAAATATACATGATCTCGAGTACAG 1530

```

```

; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-15

Alignment Scores:
Pred. No.: 2,47e-260 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
Gaps: 0
DB: 19

US-10-718-952-2 (1-510) x US-10-718-952-15 (1-1533)
Qy      1 MetPheIleGluAsnPheLysValGlyCysProAsnValIleTyrThrGluThrGluIle 20
Db      1 ATGTTTCATCAGAAATTTTAAAGGTAGAGCTCTTAATGTGAATACCAAGACTGAGATT 60
Qy      21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db      61 CAGTCGTGTACAACTACGAAACCAACCACTTGTTCACGAAACAGAAATGGCACTTAT 120
Qy      41 GlnTrrIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db      121 CAGTGATGTCGAAACCCAAATCCGTCACATTCACAAATTTAAACCAACCAACCATGTTCA 180
Qy      61 LysLeuGlyValMetLeuValGlyTrrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db      181 AATTTGGGGGTATCTGTGGGTGGGTGGAACCAACGGCTCAACCTCCACCGGTGGT 240
Qy      81 ValIleAlaAsnArgGluGlyIleSerTrrAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db      241 GTTATTCCTAACAGAGGGCATTTCAATGGCTTACAAAGACAAAGTTTCAACAGCCAT 300
Qy      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db      301 TACTTTGGCTCCCTCACCAAGCTCAGCTATTCAGATTGATCTTCCAGGAGAGGAA 360
Qy      121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db      361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAAGGTAAATCCGAGACAACTATGTGTGGG 420
Qy      141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db      421 GGATGGATATCAGCAACATGAACTGGCTGATGCAATGGCCAGGCAAAAGTGTTTGAC 480
Qy      161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db      481 ATCGATTGTGCAAGACAGATTGAGGCTTACATGAAATCATGATGTTCCACTCCCGGAATC 540
Qy      181 TyrAspProAspPheIleAlaIleAsnGlnGluGlyAlaAsnValIleLysGly 200
Db      541 TAGACCCGGATTCTTCTGCTGCAACAGAGACGCTGCCAACACAGCGATTAAAGCACCC 600
Qy      201 ThrTysGlnGlnValGlnGlnIleIleTyrAspIleValAlaPheTyrGlnValThr 220
Db      601 ACAAAGCAAGACAGATTGAGCAAAATCAACAAGACATCAAGCGTTTAAAGAAAGCACCC 660

```

RESULT 11
 US-10-718-952-15
 ; Sequence 15, Application US/10718952
 ; Publication No. US2004012873A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hiltz, William

```
QY 221 LysValAspLysValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGGACAAAGTGGTCTCTGCGACTGCAACAGAGGATATGCAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAglu 260
DB 721 GTAGCGCTTAATGACCACTGAGAAATCTTGGCTGCTGAGCAAGAATAGGCTGAG 780
QY 261 IleSerProSerThrLeuTrpAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCCACCTTGATGCCATTGCTGTGATGAAATGTTCTTTCAATTAAT 840
QY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAACAACCTTTGTACAGAGGCTGATTGATCTTGCCATCGCGAGAACT 900
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTGGTGGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTTCCTGTGGGAGCTGGATACAAAGCCACATCTAATGTTACACCACTTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyLysSerLysSer 360
DB 1021 AATGATGATATGATATCTCGGCTCCACAAACCTTCGCTCCAGGAATCTCCAAGAGC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTGTTGAGATGATGTGTCAACAGCAATGCCATCTCTATGAGCTGTGTACAATCCC 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTATTAAGATGTCCTTACGTAAGGAGATGCAAGAGCCATGAT 1200
QY 401 GlyTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValIleuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCAATGGGTGGAAGAAACACCACTTTTGCACAACATAT 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAgluLeuSer 440
DB 1261 GAGGATTCCTTCTTACTCTCTCTATTAATCTTGACCTTGCTCTTCTCTGAGCTGAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGATCCAGTTTAAAGCTGAATAAGGAGAAATTCACCTCATTTCCACCAAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
DB 1381 ACCATTCTACGCTATGTACCAAGGCTCTCTGCTTCCACCGGTACACCAAGTGGAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GGATTGTCAAGAGAGCTGCAATGCTGGAATAATGAGGCTTGTGTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
DB 1501 CCAAGAAATPAACATGATTTCTCGAGTACAAAG 1530

RESULT 12
US-10-025-003-11
; Sequence 11, Application us/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
```

```
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-11

Alignment Scores:
Pred. No.: 1,32e-259 Length: 1533
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.52% Indels: 0
DB: 14 Gaps: 0

US-10-718-952-2 (1-510) x US-10-025-003-11 (1-1533)

QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTTTCATCGAATTTTAAGGTAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyTyrTyr 40
DB 61 CAGTCCGTTGACACTACGAAACCAACCACTTGTTCACGAGACGAAATGGACCTTAT 120
QY 41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATTGTCAAAACCAATCCGTCAACTACCAATTTAAACCAACACCATGTTCCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyLysAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGAGAAACAACGGCTTACCTCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTAACAGAGAGACATTTCAATGGCTTACAAAGACAAATTCACCAACCTAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
DB 301 TACTTTGGCTCCCTACCCCAAGCCTCAGCTAATTCAGATTGATCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCTTCAGAAAGTCTGCTTCCAATGTTAAATCTTGACGACATGTGTGTGGG 420
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGAGATTCAGAACTGAACCTGGCTGATGCAAGGCGCAAGGAGGATTTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTCAGAAACAGTGAAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TAGGACCCGAGATTTTCATTTGCTGCCAACAGAGGAGCGTCCACMACAGTGAATTAAAGGC 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
DB 601 ACAAGCAAGAGCAAGTTCAGAAATCATCAAAAGCATCAAGGCGTTTAAAGAACCCACC 660
QY 221 LysValAspLysValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGGACAAAGTGGTCTCTGCGACTGCAACAGAGGATATGCAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAglu 260
```

```

Db 721 GTAGGCTTATGACACATGAGAACTCTTGCTGCTGTGACGAAATGAGGCTGAG 780
Qy 261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluValProPheIleAsn 280
Db 781 ATTTCCTTCCACCTGTGATGCAATGCTGTGTGTGATAAATGTTCTTTCATTAAAT 840
Qy 281 G1ySerProGluAsnThrPheValProG1yLeuIleAspLeuAlaIleAlaAsnThr 300
Db 841 GGAAAGCCCTCAGAACTTTGTACAGAGGCTGATGATTTCCATCGCGAGAACT 900
Qy 301 LeuIleG1yG1yAspAspPheLysSerG1yG1nThrLysMetLysSerValIleuValAsp 320
Db 901 TTGATTGTTGATGATGATCTTCAAGAGTGTCAAGCCAAATGAAATCTGTTGTTGAT 960
Qy 321 PheLeuValG1yAlaG1yIleLysProThrSerIleValSerLysAsnIleGluG1yAsn 340
Db 961 TTTCTGTGGGGCTGTGATCAAGCCAACTATAGTTAGTTACAAACCATCTGGGAAAC 1020
Qy 341 AsnAspG1yMetAsnLeuSerAlaProG1nThrPheArgSerLysG1nIleSerLysSer 360
Db 1021 AATGATGATGATGATCTCTGCTCCCAAACTTCGCTCAAGGAAATCTCCAAAGAGC 1080
Qy 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrG1uProG1yG1uIlePro 380
Db 1081 AACGTTGTTGACATATGCTCAACAGCAATGCCATCTCTATAGCTGTGTGATCAATCCC 1140
Qy 381 AspIleValIleValIleLysTyrValProTyrValG1yAspSerLysArgAlaMetAsp 400
Db 1141 GACCATGTTGTTGTTAAAGTATGTGCTTACGTAGGGAGATAGCAAGAGCCATGAT 1200
Qy 401 G1uTyrThrSerG1uIlePheMetG1yG1yLysSerThrIleValLeuIleAsnThrCys 420
Db 1201 GACTACACTTCAGAAATATTCATGCGGTGAGAAAGAACCAATGTGTTGCAACACATGT 1260
Qy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuLeuAlaG1uLeuSer 440
Db 1261 GAGGATTCCTTTTACTGCTGCTCTATTAATCTTGACTGTCTCTCTGTGAGCTGAGC 1320
Qy 441 ThrArgIleGluPheLysAlaGluAsnG1yLysPheIleSerPheIleProValAla 460
Db 1321 ACTAGATTCAGCTTAAAGCTGAATAATGAGGAAATTCACATTCATCCACCCAGTTGCT 1380
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProG1yThrProValIleAsn 480
Db 1381 ACCATTCAGCTATCTGACCAAGGCTCTGCTGTCCACCGGTACACCAAGTGTGAAAT 1440
Qy 481 AlaLeuSerLysG1nArgAlaMetLeuGluAsnIleMetArgAlaCysValG1yLeuAla 500
Db 1441 GCATTCTCAAGACAGCTGCAATGCTGGAATAATTAAGAGGCTGTGTGATTTGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuG1uTyrLys 510
Db 1501 CCAGAGATTAACATGATTCGAGTACAG 1530

```

```

; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-11
Alignment Scores:
Pred. No.: 1,32e-259 Length: 1533
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.52% Indels: 0
DB: 19 Gaps: 0
US-10-718-952-2 (1-510) x US-10-718-952-11 (1-1533)
Qy 1 MetPheIleGluAsnPheLysValIleGluCysProAsnValLysTyrThrG1uThrG1uIle 20
Db 1 ATGTCATGAGAAATTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
Qy 21 GlnSerValTyrAsnTyrG1uThrThrG1uLeuValIleG1uAsnArgAsnG1yThrTyr 40
Db 61 CAGTCGCTGATACACTACGAAACCAACGAACTTGTCAAGAGAACAGAAATGSCACTAT 120
Qy 41 G1nTPIleValIlePheProLysSerValLysTyrGluPheLysThrAsnIleIleValPro 60
Db 121 CAGTGATTTGTCAAAACCAATCCGTCACATCAATTTAAACCAACCCATGTTCCA 180
Qy 61 LysLeuG1yValMetLeuValIleTyrPGLysAsnAsnG1ySerThrLeuThrG1yG1y 80
Db 181 AAATGGGGGTGATGCTTGTTGGGTGGGTGGGAAACAAGGCTTACCTCACCGGTGT 240
Qy 81 ValIleAlaAsnArgG1uG1yIleSerTPAlaThrLysAspLysIleG1nAlaAsn 100
Db 241 GTTATTGCTTACAGAGAGACATTTTCATGGGCTACAAAGACAAAGATTCAACAGCCAAAT 300
Qy 101 TyrPheG1ySerLeuThrG1nAlaSerAlaIleArgValG1ySerPheG1nG1yG1u 120
Db 301 TACTTGGCTCCCTCACCAAGCTTCAGCTATTCGAGTGTGATCTTCCAGGGAGAGAA 360
Qy 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheG1y 140
Db 361 ATCTATGCCCACTCAAGAGTGTGCTTCAATGTTAAATCTGACCAACATGTGTTGGG 420
Qy 141 G1yTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GATGGGATATACGACAACTGAACCTGCTGATGCCATGGCCAGGCAAGGTTTGAC 480
Qy 421 GATGGGATATACGACAACTGAACCTGCTGATGCCATGGCCAGGCAAGGTTTGAC 480
Qy 161 IleAspLeuG1nLysG1nLeuArgProTyrMetG1ySerMetLeuProLeuProG1yIle 180
Db 481 ATCGATTTCCAGAGAGAGTGGAGGCTTACATGAATTCATGGTTCCATCCCGGAATC 540
Qy 181 TyrAspProAspPheIleAlaIleAsnG1nG1uLysArgAlaAsnValIleLysG1y 200
Db 541 TAGAGCCCGGATTTCAATGCTGCTCCCAACCAAGAGAGCGTGCAACATGATTAAGGAGC 600
Qy 201 ThrLysG1nG1uValG1nG1nIleIleLysAspIleLysAlaPheLysG1uAlaThr 220
Db 601 ACMAAGCAAGACCAATTCAGCAATCAACAAAGACATCAAGCGTTTAAGAGAGCAC 660
Qy 221 LysValAspLysValIleValIleuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 661 AAAGTGAACAAGTGGTGTCTGCTGAGACTGCCAACACAGAGAGATGATGCAATTTGTT 720
Qy 241 ValG1yLeuAsnAspThrMetGluAsnLeuAlaIleValAspArgAsnG1uIleG1u 260
Db 721 GTAGGCTTAAATGACCAATGAGAAATCTTGGCTGTGTGACAAATAGAGGTAG 780
Qy 261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATTTCCTTCCACCTGTGATGCAATGCTGTGTGTGATAAATGTTCTTTCATTAAAT 840

```

```
QY 281 GLYSerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
| | | | |
DB 841 GGAAGCCCTCAGAACACTTTTGACAGAGGCGATTGATCTTGCCATCGCGAGAACACT 900
QY 301 LeuIleGlyValAspAspPheIysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
| | | | |
DB 901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAGACCMAAATGTAATCTGTGGTGGTAT 960
QY 321 PheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
| | | | |
DB 961 TTCTCTGTGGGCGCTGGATCAAGCCACACATCTAAGTTACACCACTCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyLysSerLysSer 360
| | | | |
DB 1021 AATGATGATGATGATCTTCGGCTCCAAACCTTCGGCTCCAGAGAAATCTCCAAAGAC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnHisPro 380
| | | | |
DB 1081 AACGTTGTTGACGATATGGTCAACAGCAATGCCATCTCTATGAGCCTGGTCAACATCCC 1140
QY 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
| | | | |
DB 1141 GACCATGTTGTTGTTATTAAGTATGCTTACGTAAGGAGATACAAAGAGCCATGAT 1200
QY 401 GlyTyrThrSerGlyIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
| | | | |
DB 1201 GAGTACACTTCAGAGATATTCATGCGGTGGAAGAACACCATTTGTTGCAACACATCT 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuValGluLeuSer 440
| | | | |
DB 1261 GAGGATTTCCCTTTTACTCTCTCTATTAATCTTGGAATGGTCTTCTCTGAGCTGAGAC 1320
QY 441 ThrArgIleGluPheLysAlaGluValAsnGluGlyLysPheHisSerPheHisProValAla 460
| | | | |
DB 1321 ACTAGATCAGATTAAAGTGAAGATGAGGAAATTCACACTTCATCCACCCAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
| | | | |
DB 1381 ACCATTCTCAGCATCTGACCAAGAGCTCTCTGTTCCACCGGATCACACAGTGGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
| | | | |
DB 1441 GCATTGTCAAGCAGCGTGCATGCGAAACATATGAGGCGCTGTGTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
| | | | |
DB 1501 CCAGAGATTAACATGATCTCCAGTACAAAG 1530

RESULT 14
US-10-424-599-12022
; Sequence 12022, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12022
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1
US-10-424-599-12022
Alignment Scores:
```

```
Pred. No.: 3,06e-252 Length: 2018
Score: 2524.00 Matches: 483
Percent Similarity: 98.24% Conservative: 18
Best Local Similarity: 94.71% Mismatches: 9
Query Match: 95.90% Indels: 0
DB: 18 Gaps: 0

US-10-718-952-2 (1-510) x US-10-424-599-12022 (1-2018)
QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
| | | | |
DB 98 ATGTTCAATGAGAGATTTCAGGTTAGAGTCTTAACGTGAAGTACACAGAGCTGAGATT 157
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
| | | | |
DB 158 CAGTCGCGTGAACAACATTAAGAAACCACTGAATTGTTCAAGAGAAACAATATATGCACTTAT 217
QY 41 GlnTrrIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
| | | | |
DB 218 CAGTGGGTGTCMACCCCAAACTGTCAAAATATGAATTTAAACCAACACCATATCCCT 277
QY 61 LysLeuGlyValMetLeuValGlyTrrPglyValAsnAsnGlySerThrLeuThrGlyGly 80
| | | | |
DB 278 AAACTAAGGGGTAAATCTGTGGGCTGGGGGGAACAATGGCTCAACCTCTCATCTGTGT 337
QY 81 ValIleAlaAsnArgGluGlyIleSerTrrPalaThrLysAspLysIleGlnGlnAlaAsn 100
| | | | |
DB 338 GTTATTGCTAACCGAGAGGATTTTCATGGCGCAACAAAGAACAAATTCACACCAAT 397
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
| | | | |
DB 398 TACTTGGGTCCCTACCCAGGCACTAGCTAACAGATGGGGTCTTTCACAGGGGAAGA 457
QY 121 IleTrrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
| | | | |
DB 458 ATATATGCTCCATTCAAGAGCTGCTTCCATGAGAACCTCGATGATCTGTGTGGG 517
QY 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
| | | | |
DB 518 GGATGGGATATCAGTAATCTTGAACTGGCTGATGCAATGGCCAGGGCCAAAGGTTGGAT 577
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
| | | | |
DB 578 ATCGACCTGCACAAAACAGTTGAGCGGTAACAGAAATCAATGTTCCACTCCCGGAATC 637
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
| | | | |
DB 638 TATGACCCGGAATTTCAATTCCTGCTCAACCAAGAGACCGTGCACATCACTAATCAAGGCG 697
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
| | | | |
DB 698 ACAAGAAAGAAACAAATTCAACAAATCATCAAAAGACATTAGGAGATTCAAAAGAAAGACT 757
QY 221 LysValAspLysValValValLeuTrrPheAlaAsnThrGluArgTyrSerAsnLeuVal 240
| | | | |
DB 758 AAGGTTGACAAAGGTGTGTCTGTGGACAGCAACAGAGGTATACGCAACATGAT 817
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaIleValAspArgAsnGluAlaGlu 260
| | | | |
DB 818 GTGGGACTTAAGACACATGGAACCTCTTGCTTCTTGACAGAAATAGGCTGAG 877
QY 261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
| | | | |
DB 878 ATTTCCCTTCAACCTTGATGCAATGTGTGTATGGAATAATGTTCTTTCATCAAT 937
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
| | | | |
DB 938 GGAAGCCCTCAGAAACACTTTGTCCAGACATCATGATGATCTGCGCATCAAAAGAAATAT 997
QY 301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
| | | | |
DB 998 TTGATTGGAGAGATGACTTTAAAGTGTCAAGACCAAAATGAATCGGTGTGGTGAAT 1057
```

QY 321 PheLeuValGlyValIleGlyIleValSerProThrSerIleValSerTyraSerHisIleuGlyValan 340
 Db 1058 TTCTTGTGAGGGCCCGCATCAAGCCATCTATAGTGAATTGACCAATTCCTGGAAAC 1117
 QY 341 AenAspGlyMetAenLeuSerAlaProGlnThrPheArgSerIleGlyIleSerIleSer 360
 Db 1118 AATGATGCGATGAACTCTCAGCCCTCAAACTTCGCGTGAAGAAATTCGAAAGAG 1177
 QY 361 AenValValAspAspMetValAenSerAenAlaIleLeuTyrgluProGlyGluHisPro 380
 Db 1178 AATGTTGTGACCATATGTGTCAACACACGCGCATCTCATAGCGGTGTGAACATCCG 1237
 QY 381 AspHisValValIleIleValTyrgluValProTyrgluValGlyAspSerIleValIleMetAsp 400
 Db 1238 GACCATGTGTGTCTTAAAGTATGTAACCTTACGTTGGGACGCAAGAGGCGCATGAT 1297
 QY 401 GlyTyrgluSerGluIlePheMetGlyGlyIleSerThrIleValLeuHisAenThrCys 420
 Db 1298 GAGTACTCTTCTGAGATATTCAATGGCGGAAAGACCAATTCGTGTGCAACACTTGT 1357
 QY 421 GluAspSerIleuLeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
 Db 1358 GAGGATTCACGTGTACTCTCTCTATTAATCTTGCACTGTGCTCTTGTGAGCTTACG 1417
 QY 441 ThrArgIleGluPheValIleGluValGluValPheHisSerPheHisProValAla 460
 Db 1418 ACTCGAATGAGTTTAAAGCTGAATGAGGAAAGTTCCACTTCATTCACCCAGTTGGG 1477
 QY 461 ThrIleLeuSerTyrgluThrIleValProLeuValProProGlyThrProValValAen 480
 Db 1478 ACCATCTCAGTTACTGACCAAGGCCCTCTGTGCCACCGATGCGCAAGGTTAAT 1537
 QY 481 AlaLeuSerIleGluAlaMetLeuGluAenIleMetArgAlaCysValGlyLeuAla 500
 Db 1538 GCACTGTCAAGAGGAGGAGCTATGCTGAAACATCTTAAGGCGCTGTGTGATTAGCT 1597
 QY 501 ProGluAenAspMetIleLeuGluTyrglyVal 510
 Db 1598 CCTGAGAACATGATCTCTGAGTACAG 1627
 RESULT 15
 US-09-938-842A-1438
 ; Sequence 1438, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPT300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1438
 ; LENGTH: 1533
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1438
 Alignment Scores:
 Pred. No.: 3,296-237 Length: 1533
 Score: 2378.00 Matches: 448
 Percent Similarity: 95.69% Conservative: 40
 Best Local Similarity: 87.84% Mismatches: 22
 Query Match: 90.35% Indels: 0

DB: 9 Gaps: 0
 US-10-718-952-2 (1-510) x US-09-938-842A-1438 (1-1533)
 QY 1 MetPheIleGluAenPheValGlyCysProAenValIleTyrgluThrGluIle 20
 Db 1 AATGTCATGAGAGGCTTCAAGATGAGAGTCCAAATGTCAATATACACAGAGATGAGATC 60
 QY 21 GlnSerValTyrgluThrGluLeuValHisGluAenArgAenGlyThrTyrglu 40
 Db 61 AATTCGTCATGATTAACCAACACAGAGTCGTCACAGAGATGTAACGTTATCATAT 120
 QY 41 GlnThrIleValIlePheProSerValIleTyrgluPheValIleValIleValIlePro 60
 Db 121 CAATGGCTGTGAAGCCAAAGACTGTCAATATGATTTCAAAACAGACACTGCTGTCGCC 180
 QY 61 ValLeuGlyValMetLeuValGlyTyrgluIleValIleValIleValIleValIleVal 80
 Db 181 AATTAAGGCTTATGCTTGTGTGGGAGGAAACAAATGATCAACATCCGCTGTGT 240
 QY 81 ValIleAlaAenArgGluGlyIleSerTyrgluAlaThrValAspValIleGlnGlnAlaAen 100
 Db 241 GTCAATTCCTAACAAGAAATATCGTGGGAGCTAAAGACAAAGATGCAACAGCTAAT 300
 QY 101 TyrgluPheGluSerIleuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
 Db 301 TACTTGTGTCCTTATCTACGCTTCATGATTCGTTGGCTTACCAATGAGTGAGAG 360
 QY 121 IleTyrgluProPheValSerIleuProMetValAenProAspAspIleValPheGly 140
 Db 361 ACTATATGCTCTTCAAGAGTCTTCTTCCAAAGGTAATCCAGAGATGTTGTGTGT 420
 QY 141 GlyTyrgluPheIleSerAenMetAenLeuAlaAenAlaMetAlaArgAlaValPheAsp 160
 Db 421 GGTGGGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 161 IleAspLeuGluIleGluAenArgProTyrgluMetGluSerMetLeuProLeuProGlyIle 180
 Db 481 ATTGATCTGACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 181 TyrgluProAspPheIleAlaAlaAenGlnGluIleValArgAlaAenValIleValGly 200
 Db 541 TATGACCCGATTTTATGCTCTCTAATCAAGGTCACGTCGTCATATGCGTATCAAGAT 600
 QY 201 ThrTyrgluGlnGluValGlnGlnIleIleValAspIleValAlaPheValGluAlaThr 220
 Db 601 ACCAAGAAAGAAACAAGTATGATCAATCAAGGATGATGATGATGATGATGATGATGAT 660
 QY 221 ValValAspValValValIleuThrAlaAenThrGluArgTyrgluSerAenVal 240
 Db 661 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 241 ValGlyLeuAenAspThrMetGluAenLeuAlaValAlaValAspArgAenGluAlaGlu 260
 Db 721 GATGACCTTAACGATACAGAGGATATATGATGATGATGATGATGATGATGATGATGAT 780
 QY 261 IleSerProSerThrIleuTyrgluAlaIleAlaCysValMetGluAenValProPheIleAen 280
 Db 781 ATCTGCGCTTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 QY 281 GlySerProGlnAenThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAenThr 300
 Db 841 GGAAGTCTCAAAATATCTTGTCCCGCTTATCGAATGATGATGATGATGATGATGAT 900
 QY 301 LeuIleGlyValAspAspPheValSerGlyGlnThrIleMetIleSerValIleValIleAsp 320
 Db 901 TTGATTTGTGGGAGATTTTCAAGAGTGTGATGATGATGATGATGATGATGATGATGAT 960
 QY 321 PheLeuValGlyValGlyIleValProThrSerIleValSerTyrgluHisIleuGlyValan 340
 Db 961 TTCTTGTGAGAGGCTGATTAAGCTTCTGATGATGATGATGATGATGATGATGATGAT 1020
 QY 341 AenAspGlyMetAenLeuSerAlaProGlnThrPheArgSerIleGlyIleSerIleSer 360

```
Db      1021 AACGACGGGATGACCTTCAGCGCCAGACTTCCGCTTAAGAGATTCCAAGAGC 1080
Qy      361 AenValValAspAspMetValasnSerAsnAlaIleuTyrluProglylunHispro 380
Db      1081 AATGTTGTAAGACGACGTGGTAAGTAAAGGCATCTCTTGAGGCTTGTAACACCTT 1140
Qy      381 AspHisValValIleuTyrlValProTyrlValGlyAspSerLysArgAlaMetAsp 400
Db      1141 GACCATGTGCTGTCTCAAGATATGCCATATGTTGGGATAGTAAAGACAAATGGAT 1200
Qy      401 GluTyrlThrSerGluIlePheMetGlyLysSerThrIleValIleuHisAsnThrCys 420
Db      1201 GAGTACACGCTGAGATATTCAATGGAGGAGGAGAACACTATAGTGTTCATTAACACTGT 1260
Qy      421 GluAspSerLeuLeuAlaIleProIleIleuAspLeuValIleuLeuAlaGluLeuSer 440
Db      1261 GAGGACTCTCTTGCGCTGCACCAATTATCTTGATCTTGTCCTCTCGCTGAACTCAGC 1320
Qy      441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      1321 ACCAGAAATCCATTAAAGCTGAAGAGAGAGGGGAAGTTCACTTTCCACCGGTAGCT 1380
Qy      461 ThrIleLeuSerTyrlleuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db      1381 ACCATCTCAGTTAACCCTCAAAAGGCACCTCTGTACCGCCGSAACACCGGTGTTAAC 1440
Qy      481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1441 GCTTGTCGAGAGCGAGCGGCAATGCTGGAATAACATTCTTAGGGCATGTGTGGGCTTGCG 1500
Qy      501 ProGluAsnAsnMetIleLeuGluTyrlLys 510
Db      1501 CCAAGAGACACATGATCATGAGTACAAG 1530
```

Search completed: June 8, 2005, 05:08:51
Job time : 3728.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:21:22 ; Search time 3731.5 seconds
(without alignments)
5202.406 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MREINPKVCEPVRVXKTEFEI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10718952/runac.06062005.173400.12856/app.query.faeta_1.1294
-DB=EST -QFMT=fastlap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.ccl -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952@cgn2.1.6628@runac.06062005.173400.12856 -ICPU=6 -ICPU=3
-NO MMAP -LANG=EDUCERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_hic: *
4: gb_esc3: *
5: gb_esc4: *
6: gb_esc5: *
7: gb_esc6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2351	89.3	1740	9	CL961092 OsiFCC005
2	2064	78.4	2360	3	AJ583520 AJ583520
3	1986.5	75.5	1828	3	CNS9A4MK
4	1857	60.7	1813	3	EX827819 Arabidops
5	1597	60.7	1813	3	AK079323
6	1565	59.5	988	7	AK005029 Mus muscu
7	1452	55.2	936	7	CK279950 EST724028
8	1430	54.3	972	7	CK279064 EST725142
9	1428	54.3	1613	3	CV135801 EST847010
					CR599696 full-len

10	1420	54.0	953	7	CO414034	CO414034	EST844419
11	1419	53.9	914	7	CK275352	CK275352	EST721430
12	1397	53.1	898	7	CO113870	CO113870	GR_Bb014
13	1374	52.2	891	7	CV290142	CV290142	acF01-10m
14	1370	52.1	867	6	CB973407	CB973407	CAB30002
15	1367	51.9	865	7	CO085839	CO085839	GR_Ba03E
16	1366	51.6	883	6	CD438045	EL01N0508	CD438045
17	1359	51.6	906	7	CK320178	CK320178	L2P08A02
18	1357	51.6	912	7	CV142333	CV142333	EST853542
19	1353	51.4	892	7	CO409215	CO409215	EST839600
20	1350	51.3	813	7	CK768601	CK768601	Gm-r1030-
21	1349	51.3	860	7	CO072025	CO072025	GR_Ba30P
22	1338	50.8	842	7	CF211061	CF211061	CAB20007
23	1337	50.8	904	7	CV148446	CV148446	EST859655
24	1333	50.6	891	7	CF212628	CF212628	CGF100065
25	1325	50.3	836	7	CF518404	CF518404	CAP0007-1
26	1316	50.0	860	7	CO082095	CO082095	GR_Ba46C
27	1309	49.7	841	7	CO075472	CO075472	GR_Ba36C
28	1304	49.5	960	7	CK274458	CK274458	EST720536
29	1282	48.7	830	7	CO107630	CO107630	GR_Bb003
30	1276.5	48.5	1466	3	CR600268	CR600268	full-len
31	1276.5	48.5	1513	3	CR606341	CR606341	full-len
32	1265	48.1	816	6	CB975215	CB975215	CAB30006
33	1265	48.1	831	7	CO132349	CO132349	GR_Bb45C
34	1262.5	48.0	834	7	CR933094	CO133094	CGF100434
35	1257	47.8	878	7	CV137015	CV137015	EST848224
36	1252	47.6	570	7	CV147272	CV147272	EST858481
37	1249	47.5	785	6	CB682609	CB682609	OSJNBF10C
38	1229	46.7	840	7	CO112984	CO112984	GR_Bb004
39	1229	46.7	843	7	CK267514	CK267514	EST713592
40	1228	46.7	795	7	CO113324	CO113324	GR_Bb013
41	1214	46.1	739	4	BG124172	BG124172	EST459818
42	1212	46.0	799	7	CF401745	CF401745	RTW1-14
43	1211	46.0	781	6	CB343821	CB343821	CA32EN000
44	1210	46.0	778	7	CO071633	CO071633	GR_Ba30B
45	1208	45.9	779	6	CD448699	CD448699	EK07D2305

ALIGNMENTS

RESULT 1
CL961092 1740 bp DNA linear GSS 21-SRP-2004
OsiFCC005463 Oryza sativa Expressed Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.

ACCESSION CL961092.1 GI:52376905
VERSION
KEYWORDS
SOURCE

ORGANISM Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Bharatoideae; Oryzaceae; Oryza.
1 (bases 1 to 1740)

REFERENCE Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.

TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis

JOURNAL
COMMENT
Contact: Chen Chen
Department of Bioinformatic
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80486676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

Location/Qualifiers
1..1740
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"

FEATURES

source

/db.xref="taxon:39946"
 /clone.lib="Oryza sativa Express Library"
 /note="Oryza sativa exon trapped genomic sequences "

ALIGNMENT SCORES:

Pred. No.: 1.45e-272 Length: 1740
 Score: 2351.00 Matches: 449
 Percent Similarity: 94.11% Conservative: 30
 Best Local Similarity: 88.21% Mismatches: 30
 Query Match: 89.32% Indels: 0
 DB: 9 Gaps: 0

US-10-718-952-2 (1-510) x CL961092 (1-1740)

```

Oy 1 MetPheIIeGIuAsnPhelysValIGIuCySProAsnValIySerThrIuThrgIuThr 20
Db 1 ATGTTCAATCGAAGCTTCGCGTGGAGAGCCCGACGTGGTACGGCGCGGAGATC 60
Oy 21 GlnSerValIyAsnIySerIuThrIuThrgIuValHIGluAsnAArgAnGIyThrTy 40
Db 61 GAGTGGAGACTACAGATGACACAGAGAGCTGTGTGCACAGAGACGACGACGCGCTCC 120
Oy 41 GlnTrpIleValIyAsnProIySerValIySlyTyrgIuPhelysThrAsnIleHISValPro 60
Db 121 CGCTGGGTGTCGCGCCCAAGTCCGTCGCTACAACTTCGGAGCACACACCGCTCC 180
Oy 61 LysLeuGIuValMetLeuValIGIyTrpGIyIAsnAsnGIySerThrIuThrgIyGIy 80
Db 181 AAGCTGGGGGTATGTCGTGGGGGTGGGGGCGGACACAGGCTCAACGCTGACGGCTGG 240
Oy 81 ValIleAlaAsnAArgGIuGIyIleSerTrpAlaThrIyAsnIyIleGIuInIleAlaAsn 100
Db 241 GTCATCGCCACAGAGGAGGATCTCATGGCGGACCAAGACAGAGTGCAGACCAAC 300
Oy 101 TyrPheGIySerIuThrgIuIleAsnAlaIleArgValIGIySerPheGIuGIyGIu 120
Db 301 TACTATGCTCACTACACCCAGGCGTCCACATCAGGCGTGGAGCTCAACACGGGAGAG 360
Oy 121 IleTyraIleProPhelysSerIuThrIuThrProMetValIAsnProAsnIleValPheGIy 140
Db 361 ATCTAGCGCCCTTCAAGAGCCTCTGCTCCATGGTGAACCTTGATGATCTTGTGTGG 420
Oy 141 GIyTrpAsnIleSerAsnMetAsnLeuAlaAsnAlaMetAlaArgAlaIyValPheAsp 160
Db 421 GCGTGGAGACTTATGACAACTGAACCTGGCTGATGCTATGACAGGCGCAAGTACTTGAC 480
Oy 161 IleAspLeuGIuLysGIuInIleuArgProTyrmMetGIuSerMetLeuProIuProGIyIle 180
Db 481 ATTGATCTGCAGAGAGAGCTCAGACCTTACATGAGATCATGTGTGCTCTCCCGGCATC 540
Oy 181 TyrAspProAsnPhelysIleAlaIleAsnGIuGIuIleuArgAlaAsnAsnValIleGIy 200
Db 541 TATGACCCCGACTTATCGCCGCCCAACAGGATCCCGGCGCAACATTCATCAAGGGC 600
Oy 201 ThrIyGIuInIleGIuValGIuInIleIleIyAsnIleIyValAPheIySGIuAlaThr 220
Db 601 ACCAAGAAGAGACAGATGAGAGCATCATCAAGACATAGGAGATTCAAGAAAAAGAC 660
Oy 221 LysValIleAspLysValIleValIleuTrpThrAlaAsnThrGIuArgTySerAsnLeuVal 240
Db 661 AAAGTGAAGAGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Oy 241 ValIGIyLeuAsnAsnThrMetGIuInIleuAlaIleValIleAsnArgInIleuGIu 260
Db 721 GTTGGGCTCAATGACCAATGAGAACTCTCTGGGTCTGTGTGTGTGTGTGTGTGTGTGT 780
Oy 261 IleSerProSerThrIuTyraIleAlaIleAlaCySValMetGIuAsnValProPheIleAsn 280
Db 781 ATATCAACATGACATGATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Oy 281 GIySerProGIuAsnThrPheValProGIyLeuIleAspLeuAlaIleAlaArgAnThr 300

```

```

Db 841 GGGAGTCTCTCAGAACACCTTTGTGCTGGGCTGATGCATCTTGTATTAGAACATCTGC 900
Oy 301 LeuIleGIyGIyAspAspPhelysSerGIyGIuInIleIyMetLysSerValIleuValAsp 320
Db 901 CTGATGTGTGTGTGTGTATTTCAAGAGTGCACAGCAAAAGATGAACTGTGTGTGTGTAT 960
Oy 321 PheLeuValGIyAlaGIyIleIyAsnProThrSerIleValISerTyraAsnIleuGIyAsn 340
Db 961 TTCTAGTTGGTGTGGAATGAATGAAGCCCACTCAATGTGCAGCTCAACACCTGGGGAAT 1020
Oy 341 AsnAspGIyMetAsnLeuSerAlaProGIuInIlePheAsnSerIyGIuIleSerIySer 360
Db 1021 AATGATGCAATGAACCTTTCGCACTTCAACATTCCTCATCAGAGATGTCCAAAGAGC 1080
Oy 361 AsnValIleAspAspMetValIAsnSerAsnAlaIleLeuTyrgIuProGIyGIuHISPro 380
Db 1081 AATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Oy 381 AsnIleValIleValIleIySerTyraIleProTyraIleGIyAspSerIyAsnAlaMetAsp 400
Db 1141 GATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Oy 401 GIuTyTrpThrSerGIuIlePheMetGIyGIyIySerThrIleValIleHISAsnThrTy 420
Db 1201 GAGTACACCTCAGAGATCTTCAATGGGGGTGAAGACACATCGTTTGCACAAACCTGT 1260
Oy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuAlaGIuLeuSer 440
Db 1261 GAGGACTACCTTCTTGGCCGCCCATCTTCTTATCTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Oy 441 ThrArgIleGIuPhelysAlaGIuAsnGIuGIyLysPheHISerPheHISerProValAla 460
Db 1321 ACCAGATTTCAGCTGAAGCCGAGGGGAGAGAAATTCATTCCTTCCTTCATCCAGTGGCT 1380
Oy 461 ThrIleLeuSerTyTrpLeuThrIyAlaAlaProLeuValProProGIyTrpProValIAsn 480
Db 1381 ACCATCTCGAGACTTACTTCCACAGGCACTTGTCTCTGTGCACACCGAGTGTGAAC 1440
Oy 481 AlaLeuSerIyGIuArgAlaMetLeuGIuAsnIleMetArgAlaCySValIGIyLeuAla 500
Db 1441 GCCCTGGCAAGACAGAGGCGAATGCTTGAGAAACATCATGAGGCTCGGTGGCTGGCC 1500
Oy 501 ProGIuAsnAsnMetIleLeuGIuTy 509
Db 1501 CCGAGAACAAACATGATCTCGAGATAC 1527

RESULT 2
AJ583520/c 2360 bp mRNA linear EST 08-JAN-2004
LOCUS AJ583520 PpProcl Physcomitrella patens cDNA clone p001015066
DEFINITION similar to putative myoinositol-1-phosphate synthase, mRNA
sequence.
ACCESSION AJ583520
VERSION AJ583520.1 GI:40781590
KEYWORDS EST.
SOURCE Physcomitrella patens
ORGANISM Physcomitrella patens
REFERENCE Bryopsida; Funariidae; Funariaceae; Physcomitrella.
AUTHORS Hohen, A., Egener, T., Lucht, J.M., Holter, H., Reinhard, C., Schween, G.
TITLE An improved and highly standardised transformation procedure allows
efficient production of single and multiple targeted gene-knockouts
in a moss, Physcomitrella patens
JOURNAL Curr. Genet. 44 (6), 339-347 (2004)
COMMENT Contact: Schween G
Plant Biotechnology
University of Freiburg
Sommerstrasse 5, D-70104 Freiburg, Germany.
FEATURES
source
1..2360
/organism="Physcomitrella patens"

```

```

/mol_type="mRNA"
/db_xref="taxon:3218"
/clone="pp001015066"
/tissue_type="protonema"
/clone_lib="pprocl1"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      1,21e-237      Length:      2360
Score:          2064.00        Matches:      396
Percent Similarity: 88.61%      Conservative: 63
Best Local Similarity: 76.45%      Mismatches:  51
Query Match:    78.42%         Indels:       8
DB:             1             Gaps:         3

```

US-10-718-952-2 (1-510) x AJ583520 (1-2360)

```

QY      1 MetPheIIeGIuaSnPheIySValGIuCyProAsnValIySrrThrGluThrGluIle 20
DB      2150 ATGTTGGTGCAGACCTTTCGTGTGAGAGCTCTCAATGTGAAATACACGAGACATGATC 2091
QY      21 GlnSerValIyTrAsnTyrGluThrThrGluLeuValHieGIuaSnArg-----AangIy 38
DB      2090 AAGTCGAGCTATGATGACCAACACCGAGTTCCTGCACGACGCGCAAGAGCGATGGC 2031
QY      39 ThrTyrGluThrIleValIySrrProIySrrValIySrrGluThrPheIySrrIleHie 58
DB      2030 AAGTATGAGGTGGTGCACCAAGCCCAAGTCCGTGCATACCAATTTCCACTCGCCCAAG 1971
QY      59 ValProIySrrLeuGIyValMetLeuValGIyTrpGIyGIyAsnAngIySrrThrLeuThr 78
DB      1970 GTGCCCTAACCTGGAGATGATGCTCTCGTGGGTGGGACACACCGGCTCACTCTCAG 1911
QY      79 GIyGIyValIleAlaSnArgGIyGIyIleSrrTrpAlaThrIySrrPheIySrrIleGln 98
DB      1910 GCGGTGATCTCGCCCAACAAAGAGTATCTCATGGGTGACCAAGATGGCGGTGACAGA 1851
QY      98 AlaAsnTyrPheGlySrrLeuThrGlnIleAsnAlaIleHrgValGIySrrPheGln 117
DB      1850 CCGCAATTACTTGAGATCCCTCACTCAACGCTCACTGGCCGGGTGGCTCTTTCAGG 1791
QY      118 --GIyGluIuIleIyrrAlaProPheIySrrLeu---PrometValAsn-Proasp 135
DB      1790 GTGCGAAGAAAGATTCATGATCTTTAAAGAACATTTCAACGATGGTCAACCCCAAC 1731
QY      136 AspIleValPheGIyGIyTrpAspIleSrrAsnMetAsnLeuAlaAspAlaMetAlaArg 155
DB      1730 GATTCGTGATCGCGCGGTGGGATATCTCAACATGACCTGGCCGATCGCATGATCGT 1671
QY      156 AlaIySrrValPheAspIleAspLeuGlnIyGlnIleuArgProTyrMetGluSrrMetLeu 175
DB      1670 GCGCCGGTCTGGACATTCATCTTCAAGAACCACTTCCTCTCTCAATGACGATATGGTA 1611
QY      176 ProLeuProGIyIleTyrAspProAspPheIleAlaAsnGlnIyGluIuArgAlaAsn 195
DB      1610 CCCATGCGCTGATCTTCAACACCCCACTTCATCGCTGCTTCAACAGAGCGCTGTGGAAC 1551
QY      196 AsnValIleIySrrIyThrIySrrGlnIyGlnIleValGlnIleIleIySrrAspIleIySrrAla 215
DB      1550 AATTCATCAAGGGTTCCAAGAAAGAGCAATGACGCGATATATCAAGACATACCGAAC 1491
QY      216 PheIySrrValIleThrIySrrValIleAspIySrrValIleValIleuThrThrAlaAsnThrGluArg 235
DB      1490 TTCAAGGCCCAACCAACCAAGGTTCACAGAGTGAAGTGTGACCGCTTCAACACAGGCGT 1431
QY      236 TyrSerAsnLeuValIyGIyLeuAsnAspThrMet-GluAsnLeuLeuAlaIaValaIas 255
DB      1430 TACAGGACGATGATCGGTCTGATGACACCCCAAGGAAATCTGTTGATTTCTTGGA 1371
QY      255 pArgAngIuAlaGluIleSrrProSrrThrLeuTyrAlaIleAlaCySrrValMetGluAs 275
DB      1370 GAAGGAGAAATGAATCTCGCCGTCGACCTTGTATCGCTTCGCTTGATCCAGAGCA 1311

```

```

QY      275 nValProPheIleAsnGIySrrProGlnAsnThrPheValProGIyLeuIleAspLeuAl 295
DB      1310 GGTTCCTTCATCAACGCGCATGCCAGAACACATTCGCTCCGTGAGTAATTAATATACGC 1251
QY      295 aIleAlaArgAsnThrLeuIleGIyGIyAsnAspPheIySrrGIyGlnThrIySrrMetIy 315
DB      1250 TGTGAAAGAACTACTTAATTTGTGTGTGATGACGCTTCAAGAGCGGTCAACAGAAATGAA 1191
QY      315 sSerValIleuValAspPheLeuValGIyAlaGIyIleIySrrProThrSerIleValSerIy 335
DB      1190 ATCCGTTCTAGTGAATCTTCTAGTTGACGCGGAATTAAGCCAACTCAATTTGTAGCTA 1131
QY      335 rAsnHieIleuGIyAsnAsnAspGIyMetAsnLeuSrrAlaProGIyThrPheArgSerIy 355
DB      1130 CAACCATTTGGGCAACAGATGAATGAATCTGTCCGCCCTCAACATTCGTTCCAA 1071
QY      355 sGIuIleSrrIySrrAsnValIaIaAspAspMetValaAsnSerAsnAlaIleuTyrGI 375
DB      1070 GGAATCTCCAGAGCAATGTAGTGAACGACATGTGTGCTAGCAATTCATTTCTATACGC 1011
QY      375 uProGIyGluHieProAspPheIyValIaIleIySrrIyValProTyrValGIyAspSe 395
DB      1010 GCCCGGTGAACATCCCGAACCGTATTTGTGATCAAGTACGTCCCTTATGTGGGTGACAG 951
QY      395 rIySrrAlaIleAspGIyTyrThrSerGluIlePheMetGIyGIyIySrrThrIleVala 415
DB      950 CAAGAGGCGCAATGAGAGATGACATTCGAGATCTTCATAGGCGGACGTAACCAATCGT 891
QY      415 lIeuHieAsnThrCySrrGIuAspSerLeuAlaIaProIleIleIleuAspLeuValle 435
DB      890 CATGCACATTAAGTGCAGAAATTTCTTACTGCGCGCGGCTCATCTTTGACTTGTGCTT 831
QY      435 uIleuAlaGluLeuSerThrArgIleGIuPheIySrrAlaGluAngIyIyAspPheIySrr 455
DB      830 GCTGGCTGAGCTCTGCACAAGATTCATCAAGAAAGACGACAGAGAAAGTTCATTTC 771
QY      455 rPheHieProValAlaThrIleLeuSerTyrIleuThrIySrrAlaProLeuValProProGI 475
DB      770 TTTCCACCGGTTGCGTACTCTGACTCACTCAACAGGCTCCCTTGTATGCCCAAG 711
QY      475 YThrProValIaIaAsnAlaLeuSerIySrrIyGlnIyAlaMetLeuGluAsnIleuMetArgAl 495
DB      710 TACCCAGTGTGTTAACGCTTGCTTACGACAGAGGCGGATGCTTACAGAAATATCATGCGCG 651
QY      495 aCySrrValGIyLeuAlaProGluAsnAsnMetIleuGIyTyrIySrr 510
DB      650 GTGCATTTGGTGTCCCGGCAATTAATGAATGATGATGATCAAG 605

```

RESULT 3
CNS004MK
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSTRP424ZE06 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).
BX827819.1 GI:42462602
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE

CNS004MK 1828 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSTRP424ZE06 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).
BX827819.1 GI:42462602
HTC; GSTR cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1828)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Queller,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Sclanoubat M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
Unpublished
2 (bases 1 to 1828)
Genoscope.
Direct Submission

RESULT 4
AKO79323 1813 bp mRNA linear HTC 03-APR-2004
LOCUS
DEFINITION
Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:9630023013 product:MYO-INOSITOL 1-PHOSPHATE
SYNTHASE A1 (BC 5.5.1.4) homolog [Homo sapiens], full insert
sequence.
AKO79323
AKO79323.1 GI:26347916
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
MEDLINE
99279253
PUBMED
10349636
REFERENCE
AUTHORS
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
MEDLINE
20493374
PUBMED
11042159
REFERENCE
AUTHORS
3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
5 The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
6 (bases 1 to 1813)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Konda, M.,
Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakamune, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shizuku, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.jp,
URL: <http://genome.gsc.riken.jp/>, Tel:81-45-503-9222,

COMMENT
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>
URL: <http://location.riken.jp/>
FEATURES
source
1.1813
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:9630023013"
/db_xref="taxon:10090"
/clone="9630023013"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
50.1723
/note="unnamed protein product; MYO-INOSITOL 1-PHOSPHATE
SYNTHASE A1 (BC 5.5.1.4) homolog [Homo sapiens]
(SPTS|ONPH2, evidence: FASTA, 88.4%id, 100%length,
match=1668)
putative"
/codon_start=1
/protein_id="BAC37607.1"
/db_xref="GI:26347917"
/translation="MEPAEELIVDSPPVVSPEETEARVEYTRRSREGVLRVQPR
ATRFPRFARQVPLRGVLMVGNGNGSTLTAAVLANRLTPWTRGRKNGVSL
TOAGVNVGLDNGREVEVPSBALPWPAPNDVDSNDISLNTAEMRBRQVYDGC
LOEQMPHESLRPRPSVYIEFFIAQNTARDNLTIPTRAOQLQIRKDIDFRSSA
GLDKVIVMTANTERFCEVPRNDTANSLATLQGLVSESTLFAVASLIEDCAF
NGSPONTVPGLASORHVFVGDDPRSGGTAKVSYLVPLISGLKTSIVSNH
LKNNGQMLSAPIQRSKEVTSVYDVMVSHNVLVAGERPDCVVIKYVPYGD
KRALDEYSELMLGGTNTLVHNTCEDSLAPIMDLVLTLCORVSCFTSDSEP
OGFHVIVLSLSTLPAALPYPRGSRYVNALFRQSCIENTFRACVCLPQNMHLHKM
EREGGIRGEVATSPLPCKKEPTATNGCTGDANHGQATPKLSTA"
1795..1800
/note="putative"
1813
/note="putative"
polYA_signal
polYA_site
ORIGIN
Alignment Scores:
Pred. No.: 2,486-181 Length: 1813
Score: 1597.00 Matches: 301
Percent Similarity: 78.17% Conservative: 93
Best Local Similarity: 59.72% Mismatches: 104
Query Match: 60.68% Indels: 6
DB: 3 Gaps: 4
US-10-718-952-2 (1-510) x AKO79323 (1-1813)
QY 8 VALGLCVSPFOASVALYSTRThrgluThrgluIleGlnSerValTYraenTYrglu 27
DB 74 GGGAGTAGCCCGACGGTGTCTACAGCCCGAACCATCGAGCCGCTACAGTACCGG 133
QY 28 ThrrThrgluLeuLeuVHIGluLeuArgAsnTYrrTYrrTYrrIleValuysProlys 47
DB 134 ACACACCGCGGTACGCCGAG-----GGCGCGGCTGCGG-----GTGACGCCAGG 181
QY 48 SerValuysTYrgluPheLysThraIleHisValProlysLeuGlyValMetLeuVal 67
DB 182 GCTACGCGTTTACCTTCGCGACGCCCGGAGGTGGCCCGACCTCGGGGTGATGTTGGTC 241
QY 68 GYTPGGLYGLYLaanAsnGlySerThrLeuThrglyValIleAlaAsnArgGluGly 87
DB 242 GCGTGGCGGGGAACAACGCGTCCACGCTGCTGCTTTCGGCCAAATCGCTGCGC 301
QY 88 IleserTPaLathrLysAspLysIleGlnGlnAlaAsnTYrrPheGlySerLeuThrgln 107

Db	302	CTAACCTTGCCACGCGCACAGCTCGCAAGAGGCAACTTATGATCGTTACCCAG	361
Qy	108	AlAsSerAlaIleArgValGly---SerPheGlnGlyGlnIleIleTyrAlaProPheLys	126
Db	362	GGGGGACACCGTGAACCTTGCGTCTGGATGAGAACGGCCGGAGGGTGTGTGCTCCCTCAGT	421
Qy	127	SerLeuLeuProMetValAsnProAspAspIleValPheGlyGlyTTPAspIleSerAsn	146
Db	422	GGCGCTGACCACTATGGTGGCCCCCAACACACTGGGTGTGATGGCTGGGAATATCTCCGCG	481
Qy	147	MetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAspIleAspLeuGlnLysGln	166
Db	482	CTGAACCTTGCCCGAGCCCATGGCGCGCCGCGAGCTTCGACACTGGCGCTCTCAGGAACAG	541
Qy	167	LeuArgProTyrMetGlnSerMetLeuProLeuProGlyIleTyrAspProAspPheIle	186
Db	542	CTGTGGGCCCACTGGAGAGCGCTGGCTCGGGGCCCTCAGCTCATCTCCGAGTTCACTC	601
Qy	187	AlaAlaAsnGlnGlnGluArgAlaAsnAsnValIleLysGlyTyrLysGlnGlnVal	206
Db	602	GCTGCCAACACAGACAGCAGCTGGCGGACAACTCATCCCTGGCACACGTCGCCAACAGTTG	661
Qy	207	GlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThrLysValAspLysValVal	226
Db	662	GAGCAAAATCCGAAAGACATTAGAAATTTCCGATCCAGTCCGGGATTTGGAATAGCTATC	721
Qy	227	ValLeuTPTPrAlaAsnThrGlnArgTyrSerAsnLeuValValGlyLeuAsnAspThr	246
Db	722	GTGCTGTGGACCGGCATATGCGAGGCGTTTGGCGAGGGTCCAGAGTCGCAGATGACACA	781
Qy	247	MetGlnAsnLeuLeuAlaValaLysArgAsnGluValGluIleSerProSerThrIleu	266
Db	782	GCAGAAACTTGCTCATATCACTATCCAGCTTGGC---CTGGAGGTGTCACCTGCACACTT	838
Qy	267	TyrAlaIleAlaCysValMetGlnAsnAlaProPheIleAsnGlySerProGlnAsnThr	286
Db	839	TTTGCCTGGCCACGATCTGGAGGACGTGGCGCTTCTCAATGATCCCAAGAACACA	898
Qy	287	PheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIleGlyLysAspAsp	306
Db	899	CTGGTACCCGGTGGCCCTGGAACTGGCTTCGCAGGCCCATGTGTTGTATGGTGGTATAC	958
Qy	307	PheLysSerGlyGlnThrLysMetLysSerValLeuValaAspPheLeuValGlyAlaGly	326
Db	959	TTCAAGTCAGGGCAGACATAAGGTCAAGCTGTCCTGGTGGACTTCTCATCGGCTCGG	1018
Qy	327	IleLysProThrSerIleValSerTyrAsnHisIleuGlyAsnAsnAspGlyMetAsnLeu	346
Db	1019	CTCAAGACCATATGCATCGTGAAGCTATTAACCACTGGGCCAACAAGAGGGCAGAACTGT	1078
Qy	347	SerAlaProGlnThrPheArgSerTyrGlnIleSerLysSerAsnValValaAspMet	366
Db	1079	TTGGACCGCTGCAAGTTCGCTCCAAAGAGGTGACAAAGACAGTGTGTGACCATGTG	1138
Qy	367	ValaAsnSerAsnAlaIleLeuTyrGlnProGlyGlnHisProAspHisValaValaIle	386
Db	1139	GTTACACAGAACCAATGTCTCTACCGCGCTGGAGAGCGGCCACAGACCACTGTGTGTATC	1198
Qy	387	LysTyrValProTyrValGlyAspSerLysArgAlaMetAspGluTyrThrSerGlnIle	406
Db	1199	AAATATGTGCTCTATGTTGGCCACAGCAAGCCTGCTTGACAGTACACTCCGACGTG	1258
Qy	407	PheMetGlyGlyLysSerThrIleValIleuHisAsnThrCysGlnAspSerLeuLeuAla	426
Db	1259	ATGCTGGGGGACAAACACTTGGTGGCTCAATATACCTGGAGAGATTCGCTCTGGCC	1318
Qy	427	AlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGluPheLys	446
Db	1319	GGGCCCATCATGCTGACCTTATGCTGCTGCACAGAGCTGTGCACGGGTACGTTCTGC	1378
Qy	447	AlaGluAsnGlnGlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyrIleu	466
Db	1379	ACAGACTCGAACCCGAGCTCAGAGGCTTCCACAGAGTGTGCTCCCTGCTTAACTTCTGTG	1438

QY	467	Thrysalap	ProLeuVal1	ProProGlyThrProValValAsnAlaLeuSerIysGlnArg	486
Db	1439	TTTAAGGCCCCGCTTGTGCCCCCGGACGCTGTACTGATGATGCCCTTCCGACAGCG	1498		
QY	487	AlaMetLeuGluIuaenIleMeCArgAlaCysValGlyLeuAlaProGluAanAanMetIle	506		
Db	1499	AGCGTATCGAGATATTTTTCAGAGGCTTGCGTGAGGCTCCCGCCGAGAACACATGCTA	1558		
QY	507	LeuGluTyrIlys	510		
Db	1559	TTAGAGCAACAAG	1570		
RESULT 5					
AK005029					
LOCUS					
DEFINITION					
AK005029					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					

Qy 367 VALaenSerAaAla1LeuTyrgluProgiygluH1sProAph1sValaVala1le 386
Db 1140 GTTCAAGACCAACCACTGCTACGGCGCTGAGAGCGGCCAGACCACTGTGTGATTC 1199
Qy 387 lvsTyrrAlProTyrrValGlyAspSerlyAargAlaMetAspGluTyrrThsSerGluLe 406
Db 1200 AAATATGTGCTATGTTGGCCACAGACCGGTGCTTGACAGACACCTCCGACGCTG 1259
Qy 407 PheMetGlyGlyLysSerThrIleValLeuH1sAenThCyGluAspSerLeuLeuAla 426
Db 1260 ATGCTGGGGGAGCAACAACCTTGCTGCTCCATTAATACCTGAGAGATTGCTCTGGGCC 1319
Qy 427 AlaProIleIleLeuAspLeuValleuLeuAlaGluLeuSerThraArgIleGluPhelys 446
Db 1320 GGGCCATCATGCTGAGCTGAGCTGCTGCTCACAGAGCTGTCTACAGCGCTGCTTCTGC 1379
Qy 447 AlaGluAenGluGlyLysPheH1sSerPheH1sProValAlaThrIleLeuSerTyrrLeu 466
Db 1380 ACAGACTCGGACCCGAGCTGACAGGCTTCCACACAGTCTGCTGCTTACGCTTCTG 1439
Qy 467 ThrIleValProLeuValProProGlyThrProValAlaAsnSerlyAsnArg 486
Db 1440 TTAAAGCCCCCGCTGTGCCCCCGGACAGCTGTATGTAATGCCCTTCTCGCCAGGCC 1499
Qy 487 AlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluAsnMetIle 506
Db 1500 AAGTGATCGAATAATTTTTCAGGGCTTGCGTGGGCTCCGCCACAGAACACATGCTA 1559
Qy 507 LeuGluTyrrLys 510
Db 1560 TTAGAGACACAG 1571

RESULT 6
CK277950 988 bp mRNA linear EST 03-AUG-2004
LOCUS EST124028 potato abiotic stress cDNA library Solanum tuberosum cDNA
DEFINITION Clone FOAE302 5' end, mRNA sequence.
ACCESSION CK277950
VERSION CK277950.1 GI:39834928
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 988)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: EST724029
Contact: Robin Buell
The Institute for Genomic Research
9713 Medical Center Dr, Rockville, MD 20850, USA
Email: potatc-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/ .
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1..988
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="FOAE302"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at

following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	7,14e-178	Length:	988
Score:	1565.00	Matches:	301
Percent Similarity:	96.05%	Conservative:	15
Best Local Similarity:	91.49%	Mismatches:	13
Query Match:	59.46%	Indels:	1
DB:	7	Gaps:	0

US-10-718-952-2 (1-510) x CK277950 (1-988)

Qy 55 ThrAenIleH1sValProLysLeuGlyValMetLeuValGlyTyrrGlyLysAsnGly 74
Db 2 ACTGATGTCATGTTCCAAA-TTAGGGGTATGCTTGTGATGGGGGAGAAACATGCT 60
Qy 75 SerThrLeuThrGlyGlyValIleAlaAsnArgGluGlyIleSerTyrrAlaThrLysAsp 94
Db 61 TCAGCTTGACCTGAGAGGTGTTATTCCTTAACAGAGAGAAATTTCATGGGCTACCAAGAT 120
Qy 95 lvsIleGluGlnAlaAsnTyrrPheGlySerLeuThrGlnAlaSerAlaIleArgValGly 114
Db 121 AAAGTGACAGAACCAATTAATCTTGCTCTTACACAGGCTCTACATTAATCGAGTTGCT 180
Qy 115 SerPheGluGlyGluGluIleTyrrAlaProPheLysSerLeuLeuProMetValAsnPro 134
Db 181 TCCTTCAATGAGAGAGAAATCATGAGCCCTTTAAGACATCACTTCCATGGCTCATATCA 240
Qy 135 AspAspIleValPheGlyGlyTyrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAla 154
Db 241 GATGATGATGCTTTGGAGGATGGACATCAGCAATGAATTTGGCAGATGCCATGGCA 300
Qy 155 ArgAlaLysValPheAspIleAspLeuGlnLysGluLeuArgProTyrrMetGluSerMet 174
Db 301 AGGGCTAAAGTTTGGACATTCATCTCAAAAAGCAGAGGCCCTCATGGAATCCATG 360
Qy 175 LeuProLeuProGlyIleTyrrAspProAspPheIleAlaAsnGluGluLysArgAla 194
Db 361 GCCCAGCTGCGCTGATATGATGACCTGATTTGCTGCAACCAAGGCTCAGGTCT 420
Qy 195 AsnAsnValIleLysGlyTyrrLysGluGlnValGlnGlnIleIleLysAspIleLys 214
Db 421 AACATGTCATTAAGGAACCAAGAAAGAACAAATGATCAAAATCATTAAGATTTTACG 480
Qy 215 AlaPheLysGluAlaThrLysValAspLysValAlaValLeuTyrrThrAlaAsnThrGlu 234
Db 481 GAGTTTAAGAGAAACAACAAGTGACAAGGTGGTGGTTTGTGACATGCTAACACTGAA 540
Qy 235 ArgTyrrSerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaVal 254
Db 541 AGATACACAGAGTGTGCTGTGGCTTATGATGATCAAGGAAACCTTTGCTTCTGTG 600
Qy 255 AspArgAsnGluAlaGluIleSerProSerThrLeuTyrrAlaIleAlaCysValMetGlu 274
Db 601 GATGAATAATGAGCTGAATAATCTCTTCCACCTTGATGCTATTGCTTGTATTCTTGAA 660

QY 275 AaNaValProPhel1eaNglySerProGlnAaenThrPhaValProGlyLeu1leaPleu 294
 DB 661 AATGCGCCCTTCATCATGGAAGCCCAAAACACTTTTGTCCAGGCTCATGATTTTG 720
 QY 295 AAl1eAlaAaXaenThrLeu1leGlyGlyAaPaPPhelYsSerGlyGlnThrYlyMet 314
 DB 721 GCATCAAGAGAACACTTGTGATGATGATTAAGAGTGTCAACCAAGATG 780
 QY 315 LySeSerValLeuValaSpheLeuValGlyAlaGlyYl1leYProThrSer1leValSer 334
 DB 781 AAGTCAAGCTGTGATTTCTTGTGGAGCTGTATTAAGCAACATCAATTTGTGAC 840
 QY 335 TyraAen1leuGlyAaNaAaPaPglyMeCaenLeuSerAlaProGlnThrPhaYsSer 354
 DB 841 TACAACATTTGGGTAAACAATGATGAATGATCTCGGCTCAGACTTTCCGCTCA 900
 QY 355 LyGly1leSer1ySeSerAaNaValaAaPaPMeValaAaSerAaNa1le1leuYr 374
 DB 901 AAGGAGATCTCGAAAAGCAATGTTGTGATGACATGTTCTTAATGCAATCCCTTAT 960
 QY 375 GluProGlyGluH1aProAph1sVal 383
 DB 961 GAGCTGAGAGACCTGACCATGTT 987

RESULT 7
 CK279064 936 bp mRNA linear EST 03-AUG-2004
 LOCUS EST125142 potato abiotic stress cDNA library Solanum tuberosum cDNA
 DEFINITION clone POAB89 5' end, mRNA sequence.
 CK279064
 ACCESSION CK279064.1 GI:39836042
 VERSION EST.
 KEYWORDS Solanum tuberosum (potato)
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 936)
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST725143
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potcdo-array@igf.org
 Clones can be requested from the University of Arizona Genomics
 Institute via http://genome.arizona.edu/orders/
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source
 1. 936
 location/Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAB8989"
 /ciseu_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_1lb="potato abiotic stress cDNA library"
 /note="Vector: PCMVSPORT5.1; Site 1: BcoRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation
 of the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	3 06e-164	Length:	936
Score:	1452.00	Matches:	285
Percent Similarity:	95.51%	Conservative:	13
Best Local Similarity:	91.35%	Mismatches:	12
Query Match:	55.17%	Indels:	2
DB:	7	Gaps:	0

US-10-718-952-2 (1-510) x CK279064 (1-936)

QY 163 LeuGlnLySeGlnLeuAaPProTyMeGluSerMeLeuProGly1leTyraP 182
 DB 3 CTGCAAAAGCAGCTGAGGCCCTACATGAAATCCATGTTCTCTCTGATCTATGAC 62
 QY 183 ProAaPPhel1eAla1aAaNgInGluAaAaAaNaVal1leYeglyThrYs 202
 DB 63 CCGTACTTCATATGAGCTAAACCAAGAGCAGCTGCACAAAGTATCAAGAACCAAG 122
 QY 203 GInGluGlnValaGln1le1leYsAeP1leYsAlaPheLySeGluA1aThrYsVal 222
 DB 123 AAAGAAACAGTTGATCAAAATTTGTAAGATATTAGAGACTTCAAGCAAGAACAGTA 182
 QY 223 AaPlySeValaValleuTriPha1aAaenThrGluAaYTySeAaenLeuValaGly 242
 DB 183 GACAAGATAGTGGTTCTATGAGACTGCCAACCAAGAGTACAGCAATGTGTTGGC 242
 QY 243 LeuAaAaPThrMeGluAaenLeuAa1aAaAaPaPArgAaNgInAaGlu1leSer 262
 DB 243 CTTAACGATACCATGAAACCTTTAGCTGTGATGATGAAGAGGCTGAATATCT 302
 QY 263 ProSeThrLeuTyra1leAlaCySeValMeGluAaNaValProPhel1eaNglySer 282
 DB 303 CTTTCTACATTCAGATCTATTTCTGTATTCGAAATGTCTTTCTCAACGAGAC 362
 QY 283 ProGlnAaenThrPhaValProGlyLeu1leAaPLeuAa1leAlaAaXaenThrLeu1le 302
 DB 363 CCTCAAAACACTTTTGTCCAGGCTCTTATGATTTGGCATTAAGAAACCTTTAATT 422
 QY 303 GlyGlyAaPaPPhelYsSerGlyGlnThrYsMeLySeSerValleuValaSpheLeu 322
 DB 423 GGTGTGTATGACTTTTAAGAGTGTCAACCAAGAAAGAGGCTGTGTGATTTCTCTT 482
 QY 323 ValGlyAlaGly1leYsPProThrSer1leYsTyraAen1leuGlyAaNaAaP 342
 DB 483 GTTGAGCTGTATTAAGCAAGTCAATAGTACTCAATCACTTGAGTAAACAATGAC 542
 QY 343 GlyMeCaenLeuSerAlaProGlnThrPhaYsSerLySeGlu1leSerYSeSerAaVal 362
 DB 543 GGAATGAATCTTTTGTCTCTCAAACTTCGGTCTTAAGAGATCTCAAAAAGTATGTT 602
 QY 363 ValaAaPaPMeValaAaSerAaNa1le1leuYrGluProGlyGluH1aProAph1s 382
 DB 603 GTTGATGACATGTTGCTTGAACCTCAATCTCTTATGATCTGGAGAGACCTTACCAT 662
 QY 383 ValVala1leYsTyraValProTyraValaGlyAaPSeLySeAaA1MeCaPgluTy 402
 DB 663 GTCTGTGATCAATATGTTCCATATTTGAGACAGCAAGAGGCAATGATGATGAC 722
 QY 403 ThrSeGlu1lePheMeGlyGlySeSerThri1leVal1leuH1aAaenThrCySeGluAaP 422
 DB 723 ATGTCAAGATTTTTCATGGGGGAAAGAGCAACATATGATTAGCAACAACCTTGTAGGAC 782

423 SerleuLeuAlaAlaProIleIleuAspLeuValIleuLeuAlaGluIleuSerThrArg 442
Db TCTCTTTGGGAGCTCCAAATTATCTTGACTTGGTCTTCTCGCTGGAAGTCACTGCGC 842
Qy 443 IlegluPheLysAlaGlu--AsnGluGlyLysPheHisSerPheHisProValAlaThrIle 462
Db 843 ATTCAGCTCAAGCTGAAAGGGAGGGAGAACTTCACCTCTTCACCTGTGTCTACTA 902
Qy 462 IeLeuSerThrLeuThrLysAlaProLeuVal 472
Db 903 TTCTGACTACCTCCACCAAGGCTCCCTGTGTA 934
RESULT 8
LOCUS CV135801
DEFINITION 972 bp mRNA linear EST 07-SEP-2004
CDS 847-910 Sequencing ESTs from loblolly pine embryos Pinus taeda
LOCUS clone RP1A11, mRNA sequence.
ACCESSION CV135801 GI:51898116
VERSION CV135801.1 GI:51898116
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 972)
Buell, C.R., Hsieh, J., and Cairney, J.
Sequencing of ESTs from loblolly pine embryonic libraries
Unpublished (2004)
Contact: C. Robin Buell
Plant Genomics Group
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: rhnell@tigr.org
This clone is available through TIGR. Please contact pinet@tigr.org
for further information.
Location/Qualifiers
1..972
/organism="Pinus taeda"
/mol_type="mRNA"
/cultivar="7-56 mother tree, open-pollinated tree from,
Lyons, Georgia, USA"
/db_xref="taxon:3352"
/clone="RP1A11"
/lab_host="E.coli DH10B-Tona"
/note="Organ: Zygotic Embryo and Megagametophyte. Somatic
Embryo; Vector: pCMV-SPORT 6.1; Site_1: Nott; Site_2:
ScorV; tissue: whole megagametophytes isolated from pine
seeds, whole embryos excised from these megagametophytes,
whole somatic embryos and suspensor tissue from tissue
culture, isolated from cell line A12. Pooled RNA from
zygotic embryos, megagametophytes, and somatic embryos was
used for library construction. Pine cones were harvested
weekly from open-pollinated 7-56 mother trees, collections
occurred from 7/01/02 until 10/15/02. Whole
megagametophytes were first isolated from pine seeds, and
whole embryos excised from these megagametophytes each was
flash frozen. Embryo development was assessed using the
system of Pullman et al (Pullman GS, Johnson S, Peter G,
Cairney J, Xu N. 2003. Loblolly pine somatic
embryogenesis: development of a maturation medium and
resulting embryo quality. Plant Cell Reports 21:747-758
(http://link.springer.de/link/service/journals/00299/conte
nts/03/00586/)). For photographs see Clavetta et al 2001.
(Clavetta VT, Morillon R, Pullman GS, Christopels M,
Cairney J. 2001. An aquaglyceroporin is abundantly
expressed early in the development and the superior and
the embryo proper of loblolly pine (Pinus taeda L.). Plant
Physiol. 127: 1556-1567
(http://www.plantphysiol.org/cgi/content/full/127/4/1556))

Alignment Scores:

Alignment Scores:
Pred. No.: 1.49e-161 Length: 972
Score: 1430.00 Matches: 282
Percent Similarity: 95.05% Conservative: 25
Best Local Similarity: 87.31% Mismatches: 16
Query Match: 54.33% Indels: 2
Gaps: 0

US-10-718-952-2 (1-510) x CV135801 (1-972)

Qy 183 ProAspPheIleAlaAlaAsnGluGluValAlaAsnValIleLysGlyThrLys 202
Db 1 CCGATTTCATAGCGGCAACACGGGTAGAGGGCAATATATGATCAAGGCTCCAG 60
Qy 203 GlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThrLysVal 222
Db 61 AAAGGCAACCTGCAGAAATATTCAGAGATTCAGAGATTTTAAATGAAACAAAGGTG 120
Qy 223 AspLysValIleValIleuThrAlaAsnThrGluArgLysSerAsnLysValValGly 242
Db 121 GACAAAGTATGCTGACTGTGACTGTCTAATCCGAAACGATACAGGACGTTATCGTGGGT 180
Qy 243 LeuAsnAspThrMetGluAsnLeuLeuAlaAlaValAspArgAsnGluAlaGluLysSer 262
Db 181 TTTAAATGACAGAGAGAAACCTGCTTCTCTGAAAGAAATGATCGGAAATCTCG 240
Qy 263 ProSerThrLeuThrAlaAlaCysValMetGluAsnValProPheIleAsnGlySer 282
Db 241 CCATGCACTCTCTAGCGGCTTGCGCTTCAGAGAAAGCTTCCTTCAATAAAGCGCAT 300
Qy 283 ProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIle 302
Db 301 CCTCAAAACACCTTGCTTCCAGATTCGATTCATTAGCAATTCACAAAGACGATTCATC 360
Qy 303 GlyLysAspAspPheLysSerGlyLysThrLysMetLysSerValIleValAspPheLeu 322
Db 361 GCAGAGATGACTTCMAAGCGGCGACAAAGAGAAAGTGAGTCTGTGATTTCTTA 420
Qy 323 ValGlyAlaGlyIleLysProThrSerIleValSerTyAsnHisLeuGlyAsnAsnAsp 342
Db 421 GTGCGTCCCGGCTTAAAGCACTTCGATTCGACTGCATACACTCGGTACACAGCAG 480
Qy 343 GlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSerAsnVal 362
Db 481 GGGATGAATCTGTGACGCTCCACAGCTTCGGGTCAAGAAATTCACAAAGCATATGC 540
Qy 363 ValAspAspMetValAsnSerAsnAlaIleLeuTyArgLysProGlyGlnHisProAspHis 382
Db 541 GTGAGCAGATGTTGCCAGCAACGCAATCTTTACGAAACAGGGGAAACCCACAGACAT 600
Qy 383 ValValIleIleLysTyValProTyValGlyAspSerLysArgAlaMetAspGluTy 402
Db 601 GTGCGTCGATCAATATGATCTTACGTTGAGACACAGAGCCATGAGCAGATAC 660
Qy 403 ThrSerGluIlePheMetGlyLysSerThrIleValLeuHisAsnThrCysGluAsp 422
Db 661 AGTCGCGAGATTTATGAGAGAAACACCTTCATTCACCAATACATGTGAAGAT 720
Qy 423 SerleuLeuAlaAlaProIleIleuAspLeuValIleuLeuAlaGluIleuSerThrArg 442
Db 721 TCGCTTCTTGACGCTCTCTGATCTTAACTTGAATCTGCTGCGGAACTCTGACAGG 780
Qy 443 IlegluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIle 462
Db 781 ATTCACCTTAAAGAGAAAGC-GAGGGGAATTCATTTCTTCACCGGTAGCAGCAT 839
Qy 463 LeuSerTyLeuThrLysAlaProLeuValProProGlyThrProValIleAsnAlaLeu 482
Db 840 TTAAGCTACCTCCACCAAGCGGCACTGGTACCAACAGGACGCGCTGTGTTAACCCCTG 899
Qy 483 SerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGlu 502
Db 900 GCGAAGCAGAGGCGATGCTGAGAC-ATCATGAGGGCGCTGCTGCGCTCCCTCCGAA 958

QY 503 AapAmMet 505
Db 959 AACACATG 967

RESULT 9
LOCUS CRS99696
DEFINITION full-length cDNA clone CSOD1009P03 of Placenta Cot 25-normalized
of Homo sapiens (human).

ACCESSION CRS99696
VERSION CRS99696.1 GI:50480503
KEYWORDS HTC; cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1613)
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 1613)
Genoscope.

REFERENCE
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of invitrogen.
FEATURES
Location/Qualifiers
1..1613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1009P03"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.: 5,72e-161 Length: 1613
Score: 1428.00 Matches: 274
Percent Similarity: 70.55% Conservative: 83
Best Local Similarity: 54.15% Mismatches: 89
Query Match: 54.26% Indels: 60
Gaps: 4

US-10-718-952-2 (1-510) x CRS99696 (1-1613)

QY 6 PheLVaVGluCyEtrAaSnValVstYrThrGluThGluIleGlnSerValTYrAa 25
Db 56 TTCTTCGTCGAAGCGCGAGCGTCTACGGCCCGAGCCATCCAGCGCAATCGAG 115
QY 26 TTTGluThrGluLeuValHleGluAaAraGaaGlyThrTyGlnTrpIleValVys 45
Db 116 TACCGAGCAACCGCGCTCAGCCCGAG----- 142
QY 46 ProlVseValVstYrGluPheVstHraAmlHleValProlVleGluValMet 65
Db 142 ----- 142
QY 66 LeuValGlyTrpGlyGlyAaAaengVserThrLeuThrGlyGlyValIleAlaAaAraG 85
Db 143 -----GTTGGCGCT----- 151
QY 86 GluGlyIleSerTrpAlaThrVsaAryVsiIleGlnAlaAaAaTrpPheGlySerLeu 105
Db ----- 151

Db 152 -----CTCAAGAGGCGCAACTACTAGCGCTGCTG 161
QY 106 ThrGlnAlaSerAlaIleAryValGly----SerPheGlnGlyGluGluIleTYrAlaPro 124
Db 182 ACTCAGCGCGGCAACCTGAGCCCTGAGCCGAGCGGCGAGGAGGTTGTCGTAACC 241
QY 125 PheVstSerLeuLeuProMetValAaenProAaApyIleValPheGlyGlyTPAaPile 144
Db 242 TTCAAGCGGAGTGTGCTGAGTGGCGGCTGCAAGCACTGCTGTTCATAGGAGTGGACATC 301
QY 145 SerAaMetAaenValAaAaAlaAaAlaAaAaAlaAaAaAlaAaAaAlaAaAaAlaAa 164
Db 302 TCTGCTGAGAACCTGAGCCGAGGAGTACCGCGCGGCAAGTGTGTGACTGGGGCTGACG 361
QY 165 LysGlnLeuAraGProTYrMetGusSerMetLeuProLeuProGlyIleTYrAaProAaP 184
Db 362 GAGCAACTGTGGCGGCAATGAGGAGCCGCGGCGGCGGCTTGTGTTTCAATCCCGAA 421
QY 185 PheIleAlaAlaAaengGlnGluAraGAlaAaAaAaValIleLysGlyThrVstGlnGlu 204
Db 422 TTCAATGCGGCAACCAAGAGCGCGCGGAGCAACCTCAATCCAGGCTGCGCGAG 461
QY 205 GlnValGlnGlnIleIleLysApyIleValAaPheLysGluValAaThrVsaAaApyVys 224
Db 482 CAGCTGAGCAAGATCCGAGGAGCAATCCAGACTTCGCTGACCGCGGCGCTGACAA 541
QY 225 ValValValLeuTrpThrAlaAaThrGluAraGTYrSerAaenValValGlyLeuAaA 244
Db 542 GTCAATGCTGTGAGCGGCGGCAACGAGCGCTTGTGAGGTGATTCAGGCTCAAC 601
QY 245 AaPThrMetGluAaenLeuValAlaValAaAraAaGlnAlaGluIleSerProSer 264
Db 602 GACACAGCGGAGAACTGTGTGCGCACCACTTACCTGCTTCTCAATGGTCTCCGAC 658
QY 265 ThrLeuTrpAlaIleAlaCyAaVstMetGluAaenValProPheIleAaengVserProGln 284
Db 659 ACCCTTTCGCGTGCAGCATCTCGAGAGGCTGTGCTTCTCAATGGTCTCCGACG 718
QY 285 AaThrPheValProGlyLeuIleAaPheValAlaIleAaAraAaThrLeuIleGlyGly 304
Db 719 AACACCTGTGTCGCGAGACTTGTAGCTCGGTGCGACACCGGCTTTTGTGGCGGGA 778
QY 305 AaAaAaPheVstSerGlyGlnThrVstMetVstSerValLeuValAaPheLeuValGly 324
Db 779 GATGACTTCAAGTCAAGCGGCAACCAAGTCAAGTCTGTGACTTCTCATTCATTCG 838
QY 325 AlaGlyIleLysProThrSerIleValSerTYrAaHleGlnVsaAaAaApyIleMet 344
Db 839 TCCGAGCTCAAGAACATGCTCATCTGAGTTTCAACCACTGGGCAACATGGGAG 898
QY 345 AaenLeuSerAlaProGlnThrPheAraGserVsgIleSerVstSerAaenValAaAaP 364
Db 899 AACATATGCGGCGCATGAGATTCCGCTTAAGAGAGTGTCCAAAGCAACGTTGGAGC 958
QY 365 AaPmetValAaAaSerAaAaAlaIleLeuTYrGluProGlyGluIleProAaApyIleVal 384
Db 959 GACATGTGTCAAGAGCAACCAAGTCTTAAGCCCGGAGAGGCTTCAACATGCGCG 1018
QY 385 ValIleLysTYrValProTYrValGlyAaAaSerVsaAAlaAaAaApyIleTYrThrSer 404
Db 1019 GTCAATCAAGTATGTGCGGTGAGTGTGACAGCAAGCGGCGCTGATAGATACCTGG 1078
QY 405 GluIlePheMetGlyLYrVstSerThrIleValLeuHleAaAaThrCyAGluAaAaSerLeu 424
Db 1079 GAGCTGATGCTGGGCGGAAACAACAACATGTGTGCAACAACGTTGAGAGACTGGCTG 1138
QY 425 LeuAlaAlaProIleIleLeuAaAaLeuValLeuAlaGluLeuSerThrArgIleGlu 444
Db 1139 CTGGCGGCAACCATATCATGTGAGCTTGAAGCTGTGACCAAGCTGTGCGAGCGGAGAC 1198
QY 445 PheVstAlaGluAaengGlyVstPheHleAaAaSerPheHleProValAlaAaThrIleLeuSer 464
Db 1199 TTCTGCACTGAGATGAGACCGGAGCGGAGACTTTCACCCGCTGTCTGCTCTGACG 1258

QY 465 TyrLeuThrLysAlaProLeuValProProGlyThrProValValAsnAlaLeuSerLys 484
:::|||||
Db 1259 TTCCTCTTCAGAGCCGCCACTAGTCCGCCGCCAGCCCGGTGTGTAATCCGTTTCCGC 1318
QY 485 GlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAlaProGlnLysAsn 504
:::|||||
Db 1319 CAGCCGACCTGTCAGAACATCCTCAGCGCTCGTGCGCTCCGCCACAGAACAC 1378
QY 505 MetIleLeuGlnTyrLys 510
|:::|||||
Db 1379 ATGCTCTGGAACACAAA 1396
RESULT 10
CO414034 953 bp mRNA linear EST 02-JUL-2004
LOCUS CO414034
DEFINITION EST844419 Sequencing ESTs from loblolly pine embryos Pinus taeda
CDNA clone PIM182 5' end, mRNA sequence.
ACCESSION CO414034
VERSION CO414034.1 GI:49630282
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 953)
AUTHORS Buell, C.R., Hsiao, J. and Cairney, J.
TITLE Sequencing of ESTs from loblolly pine embryonic libraries
JOURNAL Unpublished (2004)
COMMENT Contact: C. Robin Buell
Plant Genomics Group
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: rbuell@tigr.org
This clone is available through TIGR. Please contact pine@tigr.org
for further information
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1..953
/organism="Pinus taeda"
/mol_type="mRNA"
/cultivar="7-56 mother tree, open-pollinated tree from,
Lyons, Georgia, USA"
/db_xref="taxon:3352"
/clone="PIM182"
/lab_host="E. coli DH10B-Tona"
/clone_lib="Sequencing ESTs from loblolly pine embryos"
/note="Organ: Zygotic Embryo and Megagametophyte. Somatic
Embryo; Vector: pCMV-SPORT 6.1; Site 1: NotI; Site 2:
EcoRV; tissue: Whole megagametophytes isolated from pine
seeds. Whole embryos excised from these megagametophytes,
whole somatic embryos and suspensor tissue from tissue
culture, isolated from cell line A12. Pooled RNA from
zygotic embryos, megagametophytes, and somatic embryos was
used for library construction. Pine cones were harvested
weekly from open-pollinated 7-56 mother trees, collections
occurred from 7/01/02 until 10/15/02. Whole
megagametophytes were first isolated from pine seeds, and
whole embryos excised from these megagametophytes each was
flash frozen. Embryo development was assessed using the
system of Pullman et al (Pullman GS, Johnson S, Peter G,
Cairney J, Xu N. 2003. Loblolly Pine somatic
embryogenesis: development of a maturation medium and
resulting embryo quality. Plant Cell Reports 21:747-758
(http://link.springer.de/link/service/journals/00299/conte
nts/03/00586/). For photographs see Clavatta et al 2001.
(Clavatta VT, Morrillon R, Pullman GS, Christeels M,
Cairney J. 2001. An aquaglyceroporin is abundantly
expressed early in the development of the suspensor and
the embryo proper of loblolly pine (Pinus taeda L.). Plant
Physiol. 127: 1556-1567
http://www.plantphysiol.org/cgi/content/full/127/4/1556/)

ORIGIN

Alignment Scores:

Pred. No.:	2,34e-160	Length:	953
Score:	1420.00	Matches:	273
Percent Similarity:	94.62%	Conservative:	26
Best Local Similarity:	86.39%	Mismatches:	17
Query Match:	53.95%	Indels:	1
DB:	7	Gaps:	0

US-10-718-952-2 (1-510) x CO414034 (1-953)

QY 115 SerPheGlnGlyGlnGlnIleTyrAlaProPheLysSerLeuLeuProMetValAsnPro 134
:::|||||
Db 7 TCCTTTAAGCGGGAAGAGATTTCATGCTCTTCAAGACCTCTTCCATGCTGATGCT 66
QY 135 AspAspIleValPheGlnGlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetLys 154
:::|||||
Db 67 AATGCAATGCTTGTGGATGGAGACAAACAAACATGAATCTAGCGGATGCCATGGCC 126
QY 155 ArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMet 174
:::|||||
Db 127 CGAGCCAGGTTCTTGACATTGACTACAGAAAGACGCTCGCGCTTACATGCAAGCATG 186
QY 175 LeuProLeuProGlyIleTyrAspProAspPheIleAlaAsnGlnGlnLysArgAla 194
|:::|||||
Db 187 ACTCCTTCTCGTATATATGACCCGGATTTCATAGCGGCAACCAAGGTGAGAGGCA 246
QY 195 AsnAsnValIleLysGlyThrLysGlnGlnValGlnGlnIleIleLysAspIleLys 214
|:::|||||
Db 247 AATAATGTGATCAAGGCTCCAGAAAGACCACTGCAAAAGATTATCCAGATATCA 306
QY 215 AlaPheLysGlnAlaThrLysValAspLysValValIleTyrThrAlaAsnThrGlu 234
|:::|||||
Db 307 GATTTTAATGCAAAACAAAGGTGACAAAGTACGTACTGTGACATGATATCCGAA 366
QY 235 ArgTyrSerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaVal 254
|:::|||||
Db 367 CGATACAGCAGCATTAATGCTGGGTTTAATGACAGAAAGCCCTACTGCTCTCG 426
QY 255 AspArgAsnGlnAlaGlnIleSerProSerThrLeuTyrAlaIleAlaCysValMetGlu 274
|:::|||||
Db 427 GAAAGAAATGAGTCCGAAATCTCGCACTGCTCTACGGCTTGCCGCTTACGAGG 486
QY 275 AsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeu 294
|:::|||||
Db 487 AACGTTCCCTCATTAAGCGAGCTCTCAAAACACTTCCTCCAGATGATGATTTA 546
QY 295 AlaIleAlaArgAsnThrLeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMet 314
|:::|||||
Db 547 GCATTTACAAAGAACAGATATCATCCAGAGATGACTTCAAGAGCGGCAACAAAGATG 606
QY 315 LysSerValLeuValAspPheLeuValGlyAlaGlyIleLysProThrSerIleValSer 334
|:::|||||
Db 607 AAGTGGGATCTGTGATTTCTTACGTGCTGCGGTCTTAAAGCCAACTTCGATTTGAGC 666
QY 335 TyrAsnIleLysGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSer 354
|:::|||||
Db 667 TACATTCACCTCGTAAACAGACGAGATGAACTGTGACGCTCCCAACATTTCCGGTCA 726
QY 355 LysGlnIleSerLysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyr 374
|:::|||||
Db 727 AAGGAATTTTCAADAGCAATGTCTGCAGCAGATGTTGCGCAAGCCATCTTTTAC 786
QY 375 GlnProGlyGlnHisProAspHisValValIleLysTyrValProTyrValIleLysP 394
|:::|||||
Db 787 GAACCAAGGGAACACCAACCATGTCTGTCACCAAGTATGATCTTACCTTGAAGC 846
QY 395 SerLysArgAlaMetAspGlnTyrThrSerGlnIlePheMetGlyGlyLysSerThrIle 414
|:::|||||
Db 847 ACCAAGAGACCATGACGATGACGTCCAGATATTTATGGAGGAAAAAC-ACCTT 905
QY 415 ValLeuHisAsnThrCysGlnAspSerLeuAlaAlaProIleIle 430

Db 906 GPRATCCACATATCATGTGAAGTTCCTTCTTGACGCTCTCTGATC 953

RESULT 11

CK275352

LOCUS CK275352 914 bp mRNA linear EST 03-AUG-2004
DEFINITION EST121430 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POADM85 5' end, mRNA sequence.

ACCESSION CK275352

VERSION CK275352.1 GI:39832330

KEYWORDS EST.

SOURCE

Solanum tuberosum (potato)
Solanum tuberosum

Organism Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.

REFERENCE

AUTHORS

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)

COMMENT

Contact: Robin Buell

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igrr.org

Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/

Seq primer: ART TMG GTG ACA CTA TAG.

FEATURES

Source

Location/Qualifiers

1..914

/organism="Solanum tuberosum"
/mol_type="mRNA"

/cultivar="Kennebec"
/db_xref="taxon:4113"

/clone="POADM85"
/tissue_type="abiotic stress created leaf and root tissue"

/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"

/note="Vector: PCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were

grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to

four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at

following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).

Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering

applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d

and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,

2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,

12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to

construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.: 2,89e-160 Length: 914
Score: 1419.00 Matches: 275
Percent Similarity: 97.95% Conservative: 11
Best Local Similarity: 94.18% Mismatches: 6
Query Match: 53.91% Indels: 0
DB: 7 Gaps: 0

US-10-718-952-2 (1-510) x CK275352 (1-914)

Qy 217 LysGUAjathLyValaAspLyValValLeuTriThrAlaAnthrGluArgTyr 236

Db 1 AAGGAAACAGCAAGATGACCAAGGTGGTGGTTTGTGACCTGCAACCTGAAGATAC 60

Qy 237 SerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArg 256

Db 61 AGCAGTGTGGCTGGTGGCTTAAATGATACCAAGGAAACCTCTTCTCTGTGGATAGA 120

Qy 257 AaNGUAlaGluIleSerProSerThrLeuTyrAlaIleCysValMetGluAsnVal 276

Db 121 AATGAGCCGAAATATCTCTTCCACCTTGATGCTATGCTTGATTTCTTGAAATGTG 180

Qy 277 ProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIle 296

Db 181 CCCTTCATCAATGGAAGACCAAAACATTTTGTGCCAGCCTCATATTTGGCCATC 240

Qy 297 AlaArgAsnThrLeuIleGlyGlyAspAspPheLeuSerGlyGlnThrLysMetLysSer 316

Db 241 AAGAGGAACACTTGTATGTGGTGTGATGACTTAAAGTGTCAACCAAGATGAAGTCA 300

Qy 317 ValLeuValAspPheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsn 336

Db 301 GTGCTGTGTAATTTCTTGTGTGAGCTGGTATTAAAGCAACATCAATGTGAGCTCAAC 360

Qy 337 HsLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlu 356

Db 361 CATTGGGTAAACATGATGAAATGAATCTGTGGCCCTCAAGCTTCCGCTAAAGAGC 420

Qy 357 ILeSerLysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluPro 376

Db 421 ATCTGAAAAGCAATGTGTGCATGACATGCTTGTAGTAATGCCATCTTATAGAGCT 480

Qy 377 GlyGluHisProAspHisValValIleLysTyrValProTyrValGlyAspSerLys 396

Db 481 GGAAGACACCTGACCATGTTGTGTGATTAAGTAACTAATGTGAGATGACCAAG 540

Qy 397 ArgAlaMetAspGlyTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeu 416

Db 541 AAGGCAATGAGATGATGACATCAAGATTTTCATGGGTGAAGAGAACCACTTATTTG 600

Qy 417 HsAsnThrCysGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeu 436

Db 601 CACAATACCTGAGAGATCTCTTGTGGCTGCCAATATATCTTGATTTGGCTCTTG 660

Qy 437 AlaGluLeuSerThrArgIleGluPheValAlaGluAsnGluGlyLysPheHisSerPhe 456

Db 661 GGTGAATCAGACACCGCATTCAGCTCAAGCTGAAGAGAGGATGAAGTTCCATCTTTC 720

Qy 457 HisProValAlaThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThr 476

Db 721 CATCCGTGGCAACCATCTCTCACTATCTTACCAAGCTCTCTGTGACCAAGATACA 780

Qy 477 ProValAlaAsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCys 496

Db 781 CCAGTGTGAATGATGCTTCAAAAGCAGAGGCAATGCTTGAACAATATATGAGGGCTGT 840

Qy 497 ValGlyLeuAlaProGluAsnAsnMetIleLeuGlu 508

Db 841 GTTGACTGTGCACCAAGAAACAATGATATTGAA 876

RESULT 12
COL113870 898 bp mRNA linear EST 16-JUN-2004

DEFINITION GR_Eb014121.r GR_Eb Gossypium raimondii cDNA clone GR_Eb014121
3', mRNA sequence.

ACCESSION COL113870

VERSION COL113870.1 GI:48812557

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 898)

AUTHORS

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Meller, C.,
Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and
Wing, R. A.

TITLE
JOURNAL

Global assembly of cotton ESTs

COMMENT

Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 014 row: 1 column: 21.

FEATURES

source

1..898
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb014121"
/tissue_type="floral"
/dev_stage="3 to +3 DEA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCWV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.: 1.29e-157 Length: 898
Score: 1397.00 Matches: 264
Percent Similarity: 95.99% Conservative: 23
Best Local Similarity: 88.29% Mismatches: 12
Query Match: 53.08% Indels: 0
DB: 7 Gaps: 0

US-10-718-952-2 (1-510) x COI13870 (1-898)

49 VALVSTYRGLuPhelysThrAsnIleHisValProlyseuGlyValMetleuValGly 68
1 GTCAAAATGAAATTCACAGCTAATACCCATGCCCTTAATTTGGGGTGAATGCTTGAGGA 60
69 TTPGlyGlyAsnAsnGlySerThrleuThrGlyValIleAlaAsnArgGlyGly 88
61 TGGGAGGAAACAATGCTTCAACCTCAGCGGTGGTATTAGCTAACAAAGAGGTATC 120
89 SerTPAlaThrLysAspLysIleGlnIleAlaAsnTyrPheGlySerleuThrGlnAla 108
121 TCTTGGGCTACTACAGACAAAGTACAAAGGCTAAATTACTTGGTTCATTGACTCAGCA 180
109 SerAlaIleArgValGlySerPheGlnGlyGlyIleTyrAlaProPheLysSerLeu 128
181 TCACAGATCCGAATGGGCTTTCACATGAGAAAGATTTATGCTTCATTAAAGGCTT 240
129 LeuProMetValAsnProAspAspIleValPheGlyGlyTyrAspIleSerAsnMetAsn 148
241 CTTCTCATGCTGAACCAATGATGATNTGTGTTTGGAGAGTGGGACATTAGTGAACATGAA 300
149 LeuAlaAspAlaMetAlaArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuArg 168
301 CTAGCTGAGCAATGGCTAGGCGCAAGGCTTTTCGACATGATGATCGCAAAAGCAACTGAGA 360
169 ProTyrMetGluSerMetLeuProLeuProGlyIleTyrAspProAspPheIleAlaAla 188
361 CCTTCATGAAATCCATGGTCCACTCCCTCGAATTCATGATTCGATTTCAATGTGCT 420
189 AsnGlnGluGluArgAlaAsnAsnValIleLysGlyThrLysGlnGlnGlnValGlnGln 208
421 AACCAAGTGAACGTGCCAATAATGCTACAAAGGGAACCAAGAAACAAGTTCAGCAG 480
209 IleIleLysAspIleLysAlaPheLysGluAlaThrLysValAspLysValValLeu 228
:::|||||

481 GTCATCAAGACATCAACGAGCTTCAAGAGAAAACAAGTGGACACGTTGTTCTACTC 540
229 TTPThrAlaAsnThrGluArgTyrSerAsnleuValIleGlyLeuAsnAspThrMetGlu 248
541 TGGACTGAAACATCGAGAGGTACAGCATGCTCATCGTGGGCTTAATGACACCGTGA 600
249 AsnleuLeuAlaAlaValAspArgAsnGluAglIleSerProSerThrleuTyrAla 268
601 AGCTTATGCTTCTTTTGGAGAGATGATACAGATTTCTCCCTCACTTGTATCT 660
269 TLeuAlaCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheVal 288
661 ATTGCTTGCTTCTTGAAGATGTTCTTTCATCAATGAGCGCCACAAACCTTTGTT 720
289 ProGlyLeuIleAspLeuAlaIleAlaArgAsnThrleuIleGlyGlyAspAspPheLys 308
721 CCAGGCTGATTGATTGGCTATTTCAAAGAACTCTGATTGGAGAGATGACTTCAAG 780
309 SerGlyGlnThrLysMetLysSerValleuValAspPheleuValGlyAlaGlyIleLys 328
781 AGTGCCAGACCAAGATGANNATCTGCTCGTGGATTTCTTGTGGGGCTGGGATCAAG 840
329 ProThrSerIleValSerTyrAsnHisleuGlyAsnAsnAspGlyMetAsnLeuSer 347
841 CCAACATGATAGTGAATGATACCACTGAGAAATTAATGATGATGAATCTGTCA 897

RESULT 13

LOCUS

CV290142 891 bp mRNA linear EST 23-SEP-2004
ao101-10ms3-d10 ao101 Asparagus officinalis cDNA clone

DEFINITION

ao101-10ms3-d10 5', mRNA sequence.

ACCESSION

CV290142

VERSION

CV290142.1 GI:52575144

KEYWORDS

EST.

SOURCE

Asparagus officinalis (garden asparagus)

ORGANISM

Asparagus officinalis (garden asparagus);
Asparagus officinalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
Asparagaceae; Asparagus.

REFERENCE

1 (bases 1 to 891)
dePamphilis, C., Carlson, J., Ma, H., Soltis, D., Soltis, P.,
Oppenheimer, D., Frohlich, M., Doyle, J., Tanksley, S., Webb, M.,
Leebens-Mack, J., Landherr, U., Ilut, D. and Wall, K.

AUTHORS

Generation of ESTs from early male inflorescences of Asparagus
officinalis

TITLE

Unpublished (2004)

JOURNAL

Unpublished (2004)

COMMENT

Contact: Claude dePamphilis or James Leebens-Mack
Mueller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131
Email: cwd3@psu.edu or jhl10@psu.edu

FEATURES

1..891
Location/Qualifiers

source

/organism="Asparagus officinalis"
/mol_type="mRNA"
/db_xref="taxon:4686"
/clone="ao101-10ms3-d10"
/tissue_type="male inflorescences"
/lab_host="SOLR"
/clone_lib="Ao101"
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
Site 2: XhoI; This is a directionally cloned,
non-normalized library. This library has been generated by
the Floral Genome Project (FGP). The Floral Genome Project
is funded by NSF's Plant Genome Research Program

ORIGIN (DBI-0115684). More information about the project can be obtained at <http://fgrp.bio.psu.edu>

Alignment Scores:
 Pred. No.: 7,72e-155 Length: 891
 Score: 1374.00 Matches: 268
 Percent Similarity: 97.58% Conservative: 14
 Best Local Similarity: 92.73% Mismatches: 7
 Query Match: 52.20% Indels: 1
 Gaps: 0

US-10-718-952-2 (1-510) x CV290142 (1-891)

```

OY 222 ValAspLeuValValLeuTTPThRAAsenThrGluabrgfyrSerLeuValVal 241
DB 4 GTrAGAAAGAGTGTGCTTGTGACTGCAAAATGAGAGATGATGCAATGATGTC 63
OY 242 GluLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 261
DB 64 GGTCTCAATGACCAATGAGAACTTGCTGCTGCTGCAAGAACGACGAGAT 123
OY 262 SerProSerThrLeuTyraAlaIleAlaCyValMetGluAsnValProPheIleAsnGly 281
DB 124 TCCCATCAACCTCTATGGTATGCTTGCGTCATGAGAACGTTCCATCATCAACGGA 183
OY 282 SerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeu 301
DB 184 ACCCCACAAACACCTTGCTGCTGCTGATGTAATGCTTCAAGAGAACAGCTCA 243
OY 302 IleGlyValAspAspPheLeuSerGlyInThrLysMetLysSerValLeuValAspPhe 321
DB 244 ATTGAGAGAGATGACTTTAAGATGACACACCAAGATGAAGCTGTGTTGATGATTC 303
OY 322 LeuValGlyAlaGlyIleAspProThrSerIleValSerTyraAsnIleGluGlyAsnAsn 341
DB 304 CTGTGAGGGGCTGATCAAGCCCACTCATGTGATGATCAACACCATCTCGGGAACAT 363
OY 342 AspGlyMetAsnLeuSerAlaProGlnInThrPheArgSerLysGluIleSerLysSerAsn 361
DB 364 GATGGATGAATCTCTGTCACACCAACCTTCGATCAAAAGAGATGATCGAAGACCAAT 423
OY 362 ValIleAspAspMetValAsnSerAsnAlaIleLeuTyrgluProGlyGluHisProAsp 381
DB 424 GTGGTGAATGACATGCTTCAACCAATGGCATCTGTATGAGCTTGCGACACCTTGAT 483
OY 382 HisValValIleLysTyraValProTyraGlyAspSerLysArgAlaMetAspGlu 401
DB 484 CATGTCAATGTTATCAAGATATGCGCTATGTTGGATAGCAAGAGGGCTATGATGAG 543
OY 402 TyrThrSerGluIlePheMetGlyGlyLysSerThrIleValIleuHisAsnThrCyGlu 421
DB 544 TACACATCTGAGATATTCATGGAGGTGCGACAGCATTTGCTGCACAAACCTGTGAG 603
OY 422 AspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSerThr 441
DB 604 GATTCGCTTTGGCTGCGCGATCACTGAGATTTGATTCATATGGCTGAGCTTACACC 663
OY 442 ArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThr 461
DB 664 AGGATTCACGCTGAGGGGAGAGGGCAAGTTCCATCTTCATCCATCAGTGGCTTACC 723
OY 462 IleLeuSerTyraLeuThrLysAlaProLeuValProProGlyThrProValValAsnAla 481
DB 724 ATCTTGAGTTACTCTCCACAGGCTCCCTTGTTCCACAGGACACACAGTGGATGATGCA 783
OY 482 LeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCyValAlaGlyLeuAlaPro 501
DB 784 CTCTCAAGACCAAGGCGATGCTGAGAACATCTCAGGCTTGATGAGCTTGCTCT 843
OY 502 GluAsnAsnMetIleLeuGluTyraLys 510
DB 844 GAGAACAC-ATGATCTCTGAGATCAAA 869

```

RESULT 14

CB973407 867 bp mRNA linear EST 01-MAY-2003
 LOCUS CAB30002_ic_Fc_D10 Cabernet Sauvignon Berry Stage I - CAB3 vitis
 DEFINITION vinifera cDNA clone CAB30002_ic_Fc_D10 5', mRNA sequence.
 ACCESSION CB973407
 VERSION CB973407.1 GI:30296613
 KEYWORDS EST:
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 867)
 Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Leslie, A., Xu, J.,
 Jones, K. and Cook, D.
 Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
 berries at various developmental stages
 Unpublished (2003)
 CONTACT: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: ACGTACCGACATATGCG.
 Location/Qualifiers
 1..867

FEATURES

source

/organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAB30002_ic_Fc_D10"
 /sex="Hermaphrodite"
 /dev_stage="Berry stage I"
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Berry Stage I - CAB3"
 /note="Organ: Berry; Vector: pDNR; Site_1: St1; Site_2:
 St11; CAB3 is a cDNA library of Vitis vinifera 'Cabernet
 Sauvignon' Clone 8 berries. Samples were collected after
 berry set from field-grown vines during stage I of berry
 growth, 17 days after full bloom. The average berry size
 was 6 millimeters. Sampled vines were located at the
 University of California, Davis, Experimental Vineyard.
 cDNAs were made by oligo-dT priming and directionally
 cloned. 5' and 3' adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGACAGTGGCATTAAGCGCGG-3' and
 5'-ATTCTAGAGCGCGGCGCCACATG-dT(30)NN-3'. Library was
 constructed using the Clontech Creator SMART Kit and
 size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Alignment Scores:

Pred. No.: 2.25e-154 Length: 867
 Score: 1370.00 Matches: 258
 Percent Similarity: 95.50% Conservative: 18
 Best Local Similarity: 89.27% Mismatches: 13
 Query Match: 52.05% Indels: 0
 Gaps: 0

US-10-718-952-2 (1-510) x CB973407 (1-867)

```

OY 19 GuileGlnSerValTyraAsnTyrgluThrTrgLeuValHisGluAsnArgAsnGly 38
DB 1 GAGATTCAATCTGTGTGACACATGAGACCAACGAGCTTGTTCATGAGAACGAGATGCG 60
OY 39 ThrTyrglnTrpIleValLysProLysSerValLysTyrgluPheLysThrAsnIleHis 58
DB 61 ACCTACCAATGGGTGTCAAGCCCAAGACTGTCAATATGAAATGATGAGACTGATGCCAT 120
OY 59 ValProLysLeuGlyValMetLeuValGlyTyrglyGlyAsnAsnGlySerThrLeuThr 78

```

```

|||||
121 GTCCCATGCTAGGGGTTATGCTTGAGGTTGGGAGGAAACAGGCTCATCTCTCACT 180
QY GYGVGVAlIleAlaAsnArgIuGIyleSerTPAlaThlyAspLysIleGInGln 98
Db 79 GYGVGVAlIleAlaAsnArgIuGIyleSerTPAlaThlyAspLysIleGInGln 98
QY 181 GGTGGGTATGATGCCATGCGAGGAGGAAATTTATGAGGCAACAGAGCAAGGTGACCA 240
Db 99 AlaAenTyPheGlySerLeuThrgInIaIaSerAlaIleArgValGIyleSerPheGInGly 118
QY 241 GGCAATTACTTTGGCTCGCTGACTCAGGATCGACCATCCGAGTTGGGCTTCAATGGA 300
Db 119 GIUGIUIleTyAlaProPheLysSerLeuLeuProMetValAsnProAspAlIleVal 138
QY 301 GAGGAGATTACGCTTCATTCAGAGCCCTCCCTATGTAACCCAGATGACATTTGTT 360
Db 139 PheGIyGIYTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIysVal 158
QY 361 TTGGGGGAGTGGGACATCGATGACATGAACCTGGGACATGATGAGCCAGGCTAAGTTC 420
Db 159 PheAspIleAspLeuGInIySGInLeuArgProTyMetGluSerMetLeuProLeuPro 178
QY 421 CTGGACATTTGATCTGACAGACAGATTGAGCCCTTACATGATGATGATGATGATGAT 480
Db 179 GIyleIeTyAspProAspPheIleAlaAlaAsnGInGInGInGInGInGInGInGIn 198
QY 481 GGAATCTATGACCTGATTTCTATGCTGCTGCAACCAATGCTGCTGCTGCTGCTGCTGCT 540
Db 199 IySGIYThrIySGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 218
QY 541 AAAGGACCAACAAAGAGAGAGGCTCCAGCATGATCATTAAGATTAATTAAGATTAAGAG 600
Db 219 AlaThrIyValAspIyValValIleuTyPThrAlaAsnThrgIuArgTySerAsn 238
QY 601 GGAACCAAGTGGATGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 239 LeuValIValGIyleuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGIn 258
QY 661 TTGATGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 259 AlaGIuIleSerProSerThrLeuTyAlaIleAlaCysValMetGluAsnValProPhe 278
QY 721 CCTGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 279 IleaAsnGlySerProGInAsnThrPheValProGIyleuIleAspLeuAlaIleAlaArg 298
QY 781 ATCATATGAGNCCCAACAGAACCTTTTGTTCAGGGCTGATTTGCTATTAAAGG 840
Db 299 AsnThrLeuIleGIyleGIyleAspAspPhe 307
QY 841 AACAGTTGATGATGATGATGATGATTTT 867

```

```

RESULT 15
CO085839 865 bp mRNA linear EST 16-JUN-2004
LOCUS GR_Ea03E19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea03E19 5',
DEFINITION mRNA sequence.
ACCESSION CO085839
VERSION CO085839.1 GI:48776473
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 865)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Scum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of cotton ESTs
Unpublished (2004)
COMMENT
TITLE JOURNAL
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona

```

```

FEATURES
Source
1..865
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea03E19"
/issue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_id="GR_Ea"
/note="Vector: pCMVSPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by Agt. More glycerol clones held in -80."
ORIGIN
Alignment Scores:
Pred. No.: 5,18e-154 Length: 865
Score: 1367.00 Matches: 262
Percent Similarity: 97.22% Conservative: 18
Best Local Similarity: 90.97% Mismatches: 8
Query Match: 51.94% Indels: 0
Gaps: 0
US-10-718-952-2 (1-510) x CO085839 (1-865)
QY 143 AspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIysValPheAspIleAsp 162
Db 2 GACATTAGTACATGAACCTAGCTGATCTCTATGCTGAGCCCAAGTTTTCGACATGAT 61
QY 163 LeuGInIySGInLeuArgProTyMetGluSerMetLeuProLeuProGIyleTyAsp 182
Db 62 CTGCAAAAGCACTGAGACCTTACATGGAATCCAGATGCCATCTCCGGAATCTACAT 121
QY 183 ProAspPheIleAlaAlaAsnGInGInGInGInGInGInGInGInGInGInGInGInGIn 202
Db 122 CCTGATTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
QY 203 GInGInGInValGInGInIleIleYsaPheIleYsaPheYsaPheYsaPheYsaPheYsa 222
Db 182 AAAGAAACAAGTTCAAGCATGATCAACAGATCAAGAGATCAAGAAAGAAACAAGGTTG 241
QY 223 AspIyValIValIleuTyPThrAlaAsnThrgIuArgTySerAsnLeuValGIyle 242
Db 242 GACAAAGTTGTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
QY 243 LeuAsnAspThrMetGluAsnLeuLeuAlaValAspAsnGInIleGIyleSer 262
Db 302 CTAAATGACACCGTGGAAAGCTTATGAGCTCTTTGGAAAGATGATCAGAGATTCT 361
QY 263 ProSerThrLeuTyAlaIleAlaCysValMetGluAsnValProPheIleaAsnGlySer 282
Db 362 CTTTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
QY 283 ProGInAsnThrPheValProGIyleuIleAspLeuAlaIleAlaArgAsnThrLeuIle 302
Db 422 CCACAAACACCTTTGTTCCAGGGTGTGATTTGGCTATTCAAGAACTGTTGATT 481
QY 303 GIyGIyAspAspPheLysSerGIyGInThrIySmetLysSerValIleuValAspPheLeu 322
Db 482 GAGGAGATGATCTTAAGATGAGGACCAACAGATGAATGTTGCTCTGTGATTTCTT 541
QY 323 ValGIyAlaGIyleIleYsaProThrSerIleValSerTyAsnHIsLeuGIyleAsnAsnAsp 342
Db 542 GTTGGGCTGGGATGACCAACATCATGATGATGATGATGATGATGATGATGATGATGAT 601
QY 343 GIyleAsnLeuSerAlaProGInThrPheArgSerIyGIyleSerIySerAsnVal 362

```

Db	602	GGCATGATCTGTGACGACCCCAACCTTCGGTCCAGAGAGATCTCAAGAGCATGTT	661
Qy	363	ValAspApmMetValAsmSerAsnAlaIleLeuTYrGluProGlyIleuHis	382
Db	662	GTGATGACATGGTTTCAAGCATGGAAATCCTGTATGACCTGGTGAAATCCTGATCAT	721
Qy	383	ValValValIleIleYrValProTYrValGlyAspSerIleValGlnIleAspGlyTYr	402
Db	722	GTGGGTCATCAATATATATGTCCATATGTGGAGACAGCAAGAGCCATGATGATGATC	781
Qy	403	ThrSerGluIlePheMetGlyGlyIleYsSerThrIleValIleuHisAsnThrCySGluAsp	422
Db	782	ACATCAGAGATATTCATGCGAGGCAAGAACACCATGTGTGCACAAACATGTAGAGAT	841
Qy	423	SerLeuLeuAlaIleProIleIle	430
Db	842	TCCCTGTGGCTGCTCCCATTAATC	865

Search completed: June 7, 2005, 22:49:15
Job time : 3757.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2005, 17:11:57 ; Search time 5315 Seconds
(without alignments)
4649.513 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MFENFKVSEPNVXKTEI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10718952/runat_06062005.173400.12846/app_query.fasta_1.1294
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPTCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=Pct -THR_MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952@cgn2.1_1.7010@runat_06062005.173400.12846 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEBUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLCG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_da:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pal:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_srs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2593	98.6	1739	8	AY038802 Glycine m
2	2593	98.6	1782	6	BD075266 Soybean P
3	2588	98.4	1533	6	BD075269 Soybean P
4	2580	98.1	1791	8	AF293970 Glycine m

5	2475	94.1	1959	8	AB009881 Nicotiana
6	2470	93.9	1931	6	AB032073 Nicotiana
7	2470	93.9	1950	8	E27176 Novel INPS
8	2468	93.8	1959	8	BT013505 Lycopersi
9	2452	93.2	1845	8	AF284065 Sesamum 1
10	2437	92.6	1954	8	BT013759 Lycopersi
11	2400	91.2	1538	8	AF120146 Trifolium
12	2400	91.2	1538	8	AF120147 Trifolium
13	2400	91.2	1538	8	AF120148 Trifolium
14	2399	91.2	1866	8	AF433879 Suaeda ma
15	2395	91.0	2053	8	U32511 Mesembryant
16	2384	90.6	1533	6	CQ080508 Sequence
17	2384	90.6	1533	6	AX056743 Sequence
18	2384	90.6	1533	8	AY143904 Arabidops
19	2384	90.6	1863	8	AY054202 Arabidops
20	2384	90.6	1864	8	AY053415 Arabidops
21	2376	90.3	1564	8	AY096554 Arabidops
22	2376	90.3	1781	6	AX054630 Sequence
23	2376	90.3	1781	8	U66307 Brassica na
24	2376	90.3	1863	8	AY065415 Arabidops
25	2373.5	90.2	1845	8	AY028259 Arabidops
26	2361	89.7	1533	8	AY323824 Xerophyta
27	2361	89.7	1890	8	ATU30250 Arabidops
28	2358	89.6	1842	8	SPM1PHSYM
29	2354	89.5	1535	6	BD073472 Regulated
30	2354	89.5	1665	6	BD073470 Regulated
31	2354	89.5	1665	8	AF056326 Zea mays
32	2354	89.5	1936	8	AB059557 Avena sat
33	2352	89.4	1914	8	AK103501 Oryza sat
34	2352	89.4	1915	8	AK058750 Oryza sat
35	2349.5	89.3	1567	8	BT001931 Arabidops
36	2349.5	89.3	1837	8	AY085931 Arabidops
37	2349.5	89.3	1870	8	AF372954 Arabidops
38	2346	89.2	1931	6	AR137808 Sequence
39	2346	89.2	1931	6	AR170167 Sequence
40	2345	89.1	1868	8	AB012107 Oryza sat
41	2338	88.9	2152	8	AF056325 Hordeum v
42	2337	88.8	1959	6	AX151207 Sequence
43	2335.5	88.5	1921	8	ATU04876 Arabidops
44	2328.5	88.5	1871	8	PV038920 Phaseolus v
45	2327.5	88.5	1978	8	CPINOIG 232632 C.paradisi

ALIGNMENTS

RESULT 1	AY038802	1739 bp	mRNA	linear	PLN 16-JUL-2001
LOCUS	AY038802				
DEFINITION	Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.				
ACCESSION	AY038802				
VERSION	AY038802.1	GI:14764465			
KEYWORDS					
SOURCE					
ORGANISM	Glycine max (soybean)				
REFERENCE	Hitc,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.				
AUTHORS	Hitc,W.D., Carlson,T.J. and Hitc,W.D.				
TITLE	Biochemical and Molecular Characterization of a Mutation that				
JOURNAL	Confers a Decreased Rafinoseaccharide and Phytic Acid Phenotype on				
REFERENCES	Soybean Seeds				
AUTHORS	Unpublished				
TITLE	2 (bases 1 to 1739)				
JOURNAL	Carlson,T.J. and Hitc,W.D.				
REFERENCES	Direct Submission				
AUTHORS	Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402,				
TITLE	Wilmington, DE 19880-0402, USA				
JOURNAL	Location/Qualifiers				
FEATURES	1..1739				
SOURCE	/organism="Glycine max"				
	/mol_type="mRNA"				

CDS

/cultivar="Wye"
/db_xref="taxon:3847"
/dev_stage="developing embryo"
11..1543
/EC_number="5.5.1.4"
/function="cyclizes glucose 6-phosphate to
L-myo-inositol-1-phosphate"
/note="mIPS-1A"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/db_xref="GI:14764466"
/translation="MEINFKVECPNVKXETETIQSYVNYETTELVHENRGTQYMWI
KPKSYKFEKTIHVPKLGVMLVGNGNNGSTLTGVALNBSGISMATKDKIQXNYF
GSLTQSAIRVGSFQGBEIVAPFKSLPMPVNPDDIVPGMDISNNLADAMARAFVD
IDIKOLARPYMSMLPLGIYDPDRIANOBERANNVKGTKPOVOOIIIDIKAFKE
ATKDVAVLVLTANTERYSNLVGINDMENLAAVDNRNBSISPTIPIALACVENV
PFINGPQNTFVPLIDLAIAENTLIGDDFKSGQTKKSVLVDLVAQAKPISIVS
YNHLGNNDGMNLAPQTFPSKEISKSNVDDVNSNALLYEGEHPDHVVVAKVYV
GDSKRAMDEYTSIEFMGKSTVLNTECEDSLAIPILIDVLIAELSTRLEFKENE
GKFHSFHPVATILSYLTKAPLVPGTPVNVNLSKQRAMLENIMRACVGLAPENNMLE
YK"

ORIGIN

Alignment Scores:
Pred. No.: 2,92e-199 Length: 1739
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.56% Indels: 0
Gaps: 0

US-10-718-952-12 (1-510) x AY038802 (1-1739)

QY 1 MetPheIIeGLuAsnPhelysValGluSerProAsnValIyetyrThrgIuThrgIuIle 20
DB 11 ATGTTTCATCGAGAATTTTAAAGGTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATT 70
QY 21 GlnSerValItyrAsnTyrgIuThrgIuLeuValHisGLuAsnArysnGlythrTyf 40
DB 71 CAGTCCGCTGACAACTACGAAACACCGAACTTGTCACGAGAACGGAATGCACTAT 130
QY 41 GlnTrpIleValIySerIySerValAsnTyrgInPhelysThrsAnthrHisValPro 60
DB 131 CAGTGGATTGCAAAACCAATCTGTCAAAATGCAATTTAAAAACCAACATTCATGCTCT 190
QY 61 LysLeuGLyValMetLeuValGlyTyrgIyGLyAsnAsnGlySerThreuthrGlyGly 80
DB 191 AATTTAGGGGTAAATGCTGTGGGTGGGTGGAACACACGCTCAACCTCAACCGGTGCT 250
QY 81 ValIleAsnArysnGLuAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 251 GTTATTGCTTAACGAGGCGATTTCATGAGCTACAAAGACAAAGATTCAACAGCCAT 310
QY 101 TyrPheGlySerLeuThrgInIaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 311 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 370
QY 121 IleTyraLpProPhelysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 371 ATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTAACTTGACACCACTATGTTTGGG 430
QY 141 GYTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 431 GGATGGGATATGACCAACATGAACCTGGCTGAATGCCATGCCAGGCAAGGTTTGGAC 490
QY 161 IleAspLeuGlnLysGlnLeuAryProTyrmecGluSerMetValProLeuProGlyIle 180
DB 491 ATCGATTTGCCAAGAGAGTGAAGCCTTTCATGAGATTCATGCTTCCATCCCGGAATC 550
QY 181 TyrAspProAspPheIleAlaIaAsnGlnGluGluAryAlaAsnAsnValIleLysGly 200
DB 551 TATGACCCCGAATTTCTGCTGCTCAACCAAGAGAGAGCTGCCAACAACGTCAAGAGGC 610

QY 201 ThrIySGInGluGlnValGInGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 611 ACAAGAGCAAGAGCAAGATTCAACAATATCATCAAGACATCAAGGGCTTTAAGAGACCAAC 670
QY 221 LysValAspLysValValIleuTrpThrIaAsnThrgIuAryGlyrSerAsnLeuVal 240
DB 671 AAAGTGACAAAGTGGTGTGACTGTGGACTGCCAACACAGAGAGGTACAGTAATTTGGTT 730
QY 241 ValGlyLeuAsnAryThrMetGluAsnLeuLeuAlaIaValAspArysnGluAlaGlu 260
DB 731 GTGGGCTTAATGACACCATGAGATATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 790
QY 261 IleserProserThreuthrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 791 ATTCTCTTCCACTTGTATGCCATTCCTGTGTGTATGAGAAATGTTCTTTATTAAT 850
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 851 GGAAGCCCTCAGAACACTTTGTGTACACAGGCTGATTGATCTTGGCCATCGCAGAGAACCT 910
QY 301 LeuIleGlyLysAspAspPheLysSerGlyInThrLysMetLysSerValLeuValAsp 320
DB 911 TTGATTGGTGAAGATGACTTCAAGAGTGTGCAGACCAAAATGAATCTGTGTGGTTAT 970
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyraAsnHisLeuGlyAsn 340
DB 971 TTCTCTTGAGGGGCGTGTATCAAGCAACATCTATAGTCAATTTCACACATCTGGGAAC 1030
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1031 AATGATGTATGAATCTTTCGGCTCCACCAACTTTCCTCCCAAGGAATCTCCCAAGAGC 1090
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrgIuProGlyGluHisPro 380
DB 1091 AACGTTGTGATGATATGTCACACAGACATGCATCTATAGAGCTGGTGAACATCA 1150
QY 381 AspHisValValIleLysTyraLpProTyraValGlyAspSerLysAryAlaMetAsp 400
DB 1151 GACCATGTTGTTTATTAAGTATGCTGCTTAACGTAGGGGACAGCAAGAGCCATGGAT 1210
QY 401 GlnTyThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAnthrCys 420
DB 1211 GAGTACACTCAGAGATATTCATGAGTGGTGAAGAGCACATTTGTTTCCACAAACATGC 1270
QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
DB 1271 GAGGATCCCTCTTAAGCTCTCTATTAATCTTGAATTGCTGCTGAGCTCAGC 1330
QY 441 ThrArgIleGlnPhelysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1331 ACTAATAATCGAGTTTAAAGCTGAATAATGAGGAAATTCACATCTATTCACACCACTGCT 1390
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1391 ACCATCTCAGCTACCTCAACCAAGGCTCTGTTTCCACCGGAGTACACAGTGTGAT 1450
QY 481 AlaLeuSerLysGlnAryAlaMetLeuGlnLeuIleMetArgAlaCysValGlyLeuAla 500
DB 1451 GCATTGTAAAGACAGCGGCAATGCTGGAACCAATAAAGAGGCTTGTGTGAGATTGCC 1510
QY 501 ProGluAsnAsnMetIleLeuGluTyTyrys 510
DB 1511 CCAAGAAATTAATGATTTCTGAGTACAAAG 1540

RESULT 2
BD075266 1782 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION
BD075266
VERSION BD075266.1 GI:22620869
KEYWORDS JP 2001519665-A/1.

SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1782)
AUTHORS Hitz,W.D. and Sebastian,S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose
JOURNAL Patent: JP 2001519665-A 1 23-OCT-2001;
BI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR13
PN JP 2001519665-A/1
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PI 08-APR-1997 US 08/835751
PR WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC saccharides
CC and phytic acid
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1782
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.:	3e-199	Length:	1782
Score:	2593.00	Matches:	502
Percent Similarity:	99.22%	Conservative:	4
Best Local Similarity:	98.43%	Mismatches:	0
Query Match:	98.56%	Indels:	0
DB:	6	Gaps:	0

US-10-718-952-12 (1-510) x BD075266 (1-1782)

QY 1 MecpheiLeGluAspPheValGluSerProAsnValIleuTyThrGluThrGluIle 20
DB 54 ATGTTTCATGAGAAATTTTAAGTTGAGTCTTAATGTAAGTACCCAGACTGAGATT 113
QY 21 GlnSerValTyraSerTygIuThrThrGluLeuValHisGluAsnArgAsnGlyThrTy 40
DB 114 CAGTCCGTGTACAACTACGAAACCAACCGAATTGTTCAAGAGAACGGAATGCACTAT 173
QY 41 GlnTPriLeValIlePheProIleSerValAsnTyrgInPheIleThrAsnThriValPro 60
DB 174 CAGTGAATGTTCAAAACCCAAATCTGTCAAAATCGAAATTTAAACCAACATTCATGTTCT 233
QY 61 LysLeuGlyValIleMetLeuValGlyTPGlyGlyAsnAsnGlySerThrIleuThrGlyGly 80
DB 224 AAATTAAGGGGTAAATGCTTGTGGGTGGGGTGAACACACGCTCAACCTTCACCGGTGT 293
QY 81 ValIleAlaAsnArgGluAspIleSerTPAlaThrIleAspIleValIleGlnIleAlaAsn 100
DB 294 GTTATTGCTTAACCGAGAGGCAATTCATGCGCTACAAAGACAAAGATTCACCAAGCCAT 353
QY 101 TyrPheGlySerLeuThrGlnIleSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 354 TACTTTGGTCTCCCTCAACCCCAAGCTTCAGGTATCCAGTTGGGTCTTCCAGGAGAGGAA 413
QY 121 IleTyrAlaProPheIleSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 414 ATCTATGCCCATTCAGAGGCTGCTCTCAATGGTTAACCTGACGACATTTGTGTTGGG 473
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
DB 474 GGATGGAGTATATCAGCAACATGAACCTGCTGATGCAATGGCAGGAGGCAAAAGGTTTGAC 533

QY 161 IleAspLeuGlnIleGlnIleuArgProTyMetGluSerMetValProIleuProGlyIle 180
DB 534 ATCGATTTCAGAAAGCAGTTGAGGCTTTCATGGAATCATGCTTCACTCCCGGAATC 593
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluIleArgAlaAsnValIleTyGly 200
DB 594 TATGACCCCGATTTCATGCTCTCCCAACCAAGAGGAGCGGCCAACACATCATCAAGGCG 653
QY 201 ThrIleGlnGluGlnValGlnGlnIleIleValAspIleValAlaPheIleGluAlaThr 220
DB 654 ACAAAGCAAGACCAAGTTCAACAAATCATCAAAAGCATCAAGCGTTTAAAGAAAGCACCC 713
QY 221 LysValAspIleValIleuThrPThrAlaAsnThrGluAspTyrSerAsnLeuVal 240
DB 714 AAAGTGAACAAAGTGGTGTACTGTGAGTGCACACAGAGAGTACGTAATTTGGTT 773
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
DB 774 GTGGGCTTTAATGACCACTGAGAAATCTTGGCTGCTGTGAGACGAATGAGCTGAG 833
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 834 ATTTCTCTTCCACCTTGATGCAATGCTTGTGTTATGAAATGTTCTTTCATTAAT 893
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 894 GGAAGCCCTCAGAAACCTTTGTACAGGCGTGAATGATCTTGCCATCCCGAGAAACACT 953
QY 301 LeuIleGlyIleAspAspPheIleSerGlyIleThrIleMetIleSerValIleuValAsp 320
DB 954 TTGATTGGTGAGAGTACCTTCAAGAGTGTGACCAAAATGAATCTGTGTGTGAT 1013
QY 321 PheLeuValGlyAlaGlyIleIleAspProIleSerIleValSerTyraSerIleuGlyAsn 340
DB 1014 TTCCTTGGGGGGCTGTATATCAAGCCCAATCATATGATCAATCAACCATCTGGGAAAC 1073
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGlnIleSerIleSer 360
DB 1074 AATGATGTATGAATCTTGGCTCACAACACTTTCCTCCAGGAAATCTCCAAAGAC 1133
QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrgIleProGlyGluIleAsp 380
DB 1134 AACGTTGTTGAGATATGTCACAGCAATGCATCTCTTAAGACCTGTGAATCTCA 1193
QY 381 AspHisValValIleIleTyraIleProTyraIleGlyAspSerIleArgAlaMetAsp 400
DB 1194 GACCATGTTGTGTATTAAGTATGTGCTTTCATGAGGAGACAGCAAGAGACCATGTGAT 1253
QY 401 GlyTyrrThrSerGlnIlePheMetGlyIleIleAsnThrIleValIleuHisAsnThrCys 420
DB 1254 GAGTACACTTCAGAAATATTCATGGGTGAAAGACACCATTTGTTGCACAAACATGTC 1313
QY 421 GluAspSerIleuLeuAlaIleProIleIleuAspLeuValIleuLeuAlaGluLeuSer 440
DB 1314 CAGGATTCCTCTTACGTGCTCTTATCTTGGACTTGGCTCTTCTGCTAGCTCAGC 1373
QY 441 ThrArgIleGlnPheIleAlaGluAsnGluGlyIlePheHisSerPheHisProValAla 460
DB 1374 ACTGAATTCAGATTAAAGCTGAATAATGAGGAAATTCATCACTTCCACCAAGTTGCT 1433
QY 461 ThrIleLeuSerTyrrLeuThrIleValAlaProLeuValProProGlyThrProValValAsn 480
DB 1434 ACCATCTTCAGTACCTCAACCAAGGCTCTGTGTTCCACCGGTACACCAATGTGTGAT 1493
QY 481 AlaLeuSerIleGlnArgAlaMetLeuGluAsnIleMetCysArgAlaCysValGlyLeuAla 500
DB 1494 GCATTGTCAAAGACGCGTCAATGCTGAAAAATGAAGAGGCTTGTGTTGATTGGCC 1553
QY 501 ProGluAsnAsnMetIleLeuGlnTyrrIle 510
DB 1554 CCAGAGAAATTAATGATATCTCGAGTACAG 1583

RESULT 3

BD075269 1533 bp DNA linear PAT 27-AUG-2002
LOCUS BD075269 Soybean plant producing seeds with reduced levels of raffinose
DEFINITION saccharides and phytic acid.
ACCESSION BD075269
VERSION BD075269.1 GI:22620872
KEYWORDS JP 2001519665-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1533)
AUTHORS Hitz,W.D. and Sepasatian,S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose
JOURNAL saccharides and phytic acid
PATENT: JP 2001519665-A 4 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR3
PN JP 2001519665-A/4
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ,SCOTT ANTHONY SEBASTIAN
PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC and phytic acid
FH Key Location/Qualifiers
FT CDS 1..1533.
LOCATION/Qualifiers
1..1533
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.:	6,31e-199	Length:	1533
Score:	2588.00	Matches:	501
Percent Similarity:	99.02%	Conservative:	4
Best Local Similarity:	98.24%	Mismatches:	5
Query Match:	98.37%	Indels:	0
DB:	6	Gaps:	0

US-10-718-952-12 (1-510) x BD075269 (1-1533)

QY 1 MetPheileGluAsnPhelysValGluSerProAsnValIysThrGluThrGluile 20
DB 1 ATGTTTCATCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCGAGACTGAGATT 60
QY 21 GlnSerValIysAsnThrGluThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCCGCTGTAACAACGAAACCAACCGAATGTTTCAAGAAACAGGATGGCACTAT 120
QY 41 GlnTPriLeValIysProIysSerValAsnTyrGlnPhelysThrAsnThrHisValPro 60
DB 121 CAGTGATTTGCAAAACCAAAATCTGTCAAAATAGCAATTTAAACCAACATCTCATCTTCT 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrIleuThrGlyGly 80
DB 181 AAATTAGGGGTATATGTTGTGGGTTGGGGTGGAAACACGCTCAACCTCACCGGTGTT 240
QY 81 ValIleAlaAsnArgGluAspIleSerTTPAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTTGCTAACCGAGAGGCAATTTTCATGGGCTACAAAGACAAGATTCACAAAGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGluGlu 120
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGTCCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140

DB 361 ATCTATGCCCATTCAGAGAGCTGCTTCCATGTTAACTCCCTGACGACATGTGTTGGG 420
QY 141 GlnTyrPheAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGAATGGATATCAGCAACATTAACCTGCTATATGCCATGCGCAGAGCAAGAGTGTTCAC 480
QY 161 IleAspLeuGlnIysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTTGCAAGACAGATGAGGCTTACATGATCATCATGCTTCCACTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnIysGluValGlnAlaAsnValIleIleGly 200
DB 541 TATGACCCGAGTTTCATTCATTCGCAACCAAGAGCGCTGCCAACAGTCATCAAGGCG 600
QY 201 ThrLysGlnIysGlnIysValGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACNAGCAAGAGCAGATTCACAAATCTCAAAAGACATCAAGGCGTTTAAAGAACCCAC 660
QY 221 LysValAspLysValValIleuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGACAAGTGGTGTGTACTGTGACTGCCAACACAGAGAGGTACAGTAATTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTAATGACACCATGAGAGATCTTGGCTGTCGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCCTTCACCTTTATGACCATCTTGTGTGTATGAAATGTTCTTTCATTAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATGATCTTGCATGCGCAGGAACACT 900
QY 301 LeuIleGlyIysAspAspPheLysSerGlyThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTTGGTGAAGATGACTTCAGAGTGTGCAGACCAAAATGAAATCTGTGTGGTTAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
DB 961 TTCTTGGGGGGCTGGATGATCAAGCAACATCTATAGCATTCACATCACTGGGAAC 1020
QY 341 AsnAspGlyMetCysLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGTTATGAATCTTTCGGCTCCACAACTTCCGTTCCAGGAATCTCCAGAGAC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTTGATGATATGATGTCACACAGCAATGCCATCTCTATGAGCCCTGGTGAACATCCA 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTATTAATGATATGCTCTTACGTAAGGGACAGAAATAGAACCTGAT 1200
QY 401 GlnTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGAGTGAAGAAGCACCATGTTTCCACAAACATGTC 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTTCGCTGCTCTTATATCTTGGACTTGTGCTTGTGCGAGCTGACG 1320
QY 441 ThrArgIleGlnPhelysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGAAATCGATTAAAGCTGAATAATAGGAAATATCCATCTTCCACCCAGTTCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATCTTCAGTACTTACCAAGGCTCTCTGTTCACCGGGATACCAAGTGTGAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500

DB 1441 GCATTGTCAAGACGCGCATCTGTGAAAAACAAACAAATGAGGCGCTTGCTTGCAATGCC 1500

OY 501 ProGluAenAenMetIleuGluTyrIys 510
 |||||
 DB 1501 CCAGAGATAATACATGATTCTCGAGTACAAG 1530

RESULT 4

AP293970 1791 bp mRNA linear PLN 03-MAY-2001

LOCUS Glycine max myo-inositol-3-phosphate synthase (MIPS) mRNA, complete cds.

ACCESSION AF293970

VERSION AF293970.1 GI:13936690

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 1791)

AUTHORS Hegeman,C.E., Good,L.L. and Grabau,E.A.

TITLE Expression of D-myo-inositol-3-phosphate synthase in soybean. Implications for phytic acid biosynthesis

JOURNAL Plant Physiol. 125 (4), 1941-1948 (2001)

MEDLINE 21196082

PUBMED 11299373

REFERENCE 2 (bases 1 to 1791)

AUTHORS Hegeman,C.E., Good,L.L. and Grabau,E.A.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA 24060, USA

FEATURES

source location/Qualifiers

1..1791

/organism="Glycine max"

/mol_type="mRNA"

/cultivar="Williams 82"

/db_xref="taxon:3847"

/dev_stage="immature cocolyedons"

1..1791

/gene="MIPS"

62..1594

/gene="MIPS"

/note="isomerase"

/codon_start=1

/product="myo-inositol-3-phosphate synthase"

/protein_id="AAK49896.1"

/db_xref="GI:13936691"

/translation="MFIENFVCEPNKYETETIQSVYNYETTELVEHNNGTYQWITV
 KPKGVAKYFKTNHVPKLGVALVGMGGNGSTLTGCVIYNREGISNATKDKIQANVFP
 GILQASAIRVGSFQGEIYAPFSLPMVPPDDIVFGMDISNNLADMAARAKFVD
 IDLQKORPVMESMLPLGIYDPFLIANOEBERANNYIKGKROVOQIIKDKAFKE
 ATKDKVIVMTANTERYSNVGLNMTMELLAVDNRNKEISPTLIACMEANV
 PFIKGSPONTFVPGELIDAIKAKTLIGDDPKSQDTQKSVLVPLVAGIKPYSIVS
 YNLGKNDGKMLSAPOTFRSKEISKSNVDDVNSMLIVEPGHPHVVVIXKVPVY
 GDSGRADVETYSIFMGCKSTIVLANTCESLAAPIILDIVLAEISTEFAENRE
 GKFSFSPVATITLSYLTAKAPLVPGTGVNLSKRAMLEINIMACVGLPENNMILE
 YK"

ORIGIN

Alignment Scores:

Pred. No.: 3.37e-198 Length: 1791

Score: 2580.00 Matches: 500

Percent Similarity: 98.82% Conservative: 4

Best Local Similarity: 98.04% Mismatches: 6

Query Match: 98.06% Indels: 0

DB: 8 Gaps: 0

US-10-718-952-12 (1-510) x AF293970 (1-1791)

OY 1 MetPheIleGluAenPheIysValIGluSerProAenValIysIYrThrGluThrGluIle 20
 |||||

Db	62	ATGTTCAATCGAAGATTTTAAGTTGAGTGTCTAATGTGAAGTACACCGAGCTGAGAT	121
QY	21	GIInSeRvAlTYrAsnTYrGIuThrThrGIuLeuValhIsgIuAsnArgAsnGIYThTYr	40
Db	122	CAGTCCGATGTACAACTACGAAACCAACCGAACTTGTTCAACGAAACAGGAATGGCACTAT	181
QY	41	GIInTPPIleValvSPProLySeSerAlaSnTYrGIInPhelyThrAsnThHISValPro	60
Db	182	CAGTGAATGTGCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAACATCATCTTCTT	241
QY	61	LYsLeuGIYValMeTLeuValGIYTPrGIYGIYAsnAsnGIYSeThrLeuThrGIYGIY	80
Db	242	AAATTAAGGGGATTAATCTTGTTGGGTGGGGTGGAAACCAACGGCTCAACCTCACCGGTGT	301
QY	81	ValIleAlaAsnArgIuAspIleSeRTpAlaThrLYsAspLYsIleGIInAlaAsn	100
Db	302	GTTATGTGTAACCGAGAGGGCATTTCAITGGGCTACAAAGACAAAGATTCAACAAAGCAAT	361
QY	101	TYrPheGIYSeRLeuThrGIInAlaSeRAlaIleArgValGIYSePhelngIYGIuGIu	120
Db	362	TACTTGGCTCCCTCACCAAGCCCTACACTATCCAGATTGGGTCTTCCAGGAGAAAGAA	421
QY	121	IleTYrAlaPRophelySeRLeuLeuProMetValAsnPRoAspIleValPhelGIY	140
Db	422	ATCTATGCCCATTTCAAGAGCTGCTCTTCCATGGTTAACCCTGAGACATGTGTGTGG	481
QY	141	GIYTPAspIleSeRAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLYsValPheAsp	160
Db	482	GGATGGGATATCAGAACATGACCTGGCTGATGGCAITGGCAGGCGCAAGGTGTGGAC	541
QY	161	IleAspLeuGIInLYsGIInLeuArgPRoTYrMetGIuSeRMetValProLeuPRoGIYIle	180
Db	542	ATCGATTGTGCAAGACAGTGGAGGGCTTACATGAAATCAATGCTTCCACTCCCGGAAATC	601
QY	181	TYrAspPRoAspPheIleAlaAlaAsnGIInGIuAspAlaAsnAsnValIleLYsGIY	200
Db	602	TATGACCCGGAATTTATGTGTGGCCAAACCAAGAGCGTGGCCAAACAGTCACTCAAGGGC	661
QY	201	ThrLYsGIInGIuGIInValGIInIleIleLYsAspIleLYsAlaPhelySeGIuAlaThr	220
Db	662	ACAAAGCAAGAGCAAGTTCACAAATATCTCAAGACATCAAGGCGTTTAAGGAACCAAC	721
QY	221	LYsValAspLYsValValIleuTPThrAlaAsnThrGIuArgTYrSeRAsnLeuVal	240
Db	722	AAAGTGGCAAGGTGTGTACTGTGGACTGCCAAACAGAGAGGTACAGTAATTTGGTT	781
QY	241	ValGIYLeuAsnAspThrMetGIuAsnLeuLeuAlaValAspArgAsnGIuAlaGIu	260
Db	782	GTGGGCGCTTAATGACACCATGAGAGATCTTGGCTGCTGTGGACAGAAATGAGGCTGAG	841
QY	261	IleSeRPRoSeRTpIleuTYrAlaIleAlaCYsValMeTGIuAsnValProPhelIleAsn	280
Db	842	ATTTCCTCTTCACCTGTATGCACTTGCTGTGTATGAGAAATGTTCTTCAATTAAAT	901
QY	281	GIYSeRPRoGIInAsnThrPheValProGIYLeuIleAspLeuAlaIleAlaArgAsnThr	300
Db	902	GGAAGCCCTCAGAAACACTTTGTATGCACAGGGCTGATTGATCTTGGCATGGCAGAGAACT	961
QY	301	LeuIleGIYGIYAspAspPheLYsSeRGIYInThrLYsMetLYsSeRValLeuValAsp	320
Db	962	TTGATGTGTGAGATGATTTCAAGAGTGTGAGACCAAAATGAATATCTGTGTGGTTGAT	1021
QY	321	PhelLeuValGIYAlaGIYIleLYsPRoThrSeRIleValSeTYrAsnHISLeuGIYAsn	340
Db	1022	TTCTCTGTGGGGGCGTGTATCAACGCAACATCTATGCTCACTTCAACCACTGTGGGAAC	1081
QY	341	AsnAspGIYMeTAsnLeuSeRAlaPRoGIInThrPheArgSeRLYsGIuIleSeRLYsSeR	360
Db	1082	AATGATGGTATGAATCTTTCCGCTCCAAACCTTCCGTTCCAAAGAAATCTCCAAAGAC	1141
QY	361	AsnValValAspAspMetValAsnSeRAsnAlaIleLeuTYrGIuPRoGIYGIuHISPRo	380
Db	1142	AACGTGTGTATGATGTGTCAACAGCAATCCATCTTATGAGCCCTGTGAACATCCA	1201

QY 381 Asph1sValValValIleuSerTyValProTyValGlyAspSerLysArgAlaMetAsp 400
DB 1202 GACCATGTTGTTGTTATTAAGATATGCTTACGTAAGGAGACAGACAGACCATGAT 1261
QY 401 GtUTyTyrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuHisAsnThrCys 420
DB 1262 GAGTACACTTCAGATATATTCATGCGTGGAAAGACACCATGTTGTCACAAACATGCG 1321
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1322 GAGGATTCCTCTTACGCTGCTCTATTAATCTTGACCTTGCTCTTCTTGCTGAGCTCAGC 1381
QY 441 ThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1382 ACTAGATCGATTTTAAGCTGAATGAGGAAATGAGGAAATTCATCTATTCACACCATGCT 1441
QY 461 ThrIleLeuSerTyLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1442 ACCATATCTACGCTACCTCACCAAGGCTCCTCGTGTCCACCGGTAACACCGGCTGAT 1501
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1502 GCATTGTCAAGACGCGTCAATGCTGGAACATTAATGAGGCTTGTTGTTGATTGGCC 1561
QY 501 ProGluAsnAsnMetIleLeuGluTyTyLys 510
DB 1562 CCAGGAATTAACATGATTCTCGAGTAACAAG 1591
RESULT 5
AB009881 1959 bp mRNA linear PLN 13-JUL-2000
LOCUS Nicotiana tabacum mRNA for myo-inositol 1-phosphate synthase,
DEFINITION complete cds.
ACCESSION AB009881
VERSION AB009881.1 GI:8096265
KEYWORDS myo-inositol 1-phosphate synthase.
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE
AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
TITLE 1 (sites)
Screening of wound-responsive genes identifies an immediate-early
expressed gene encoding a highly charged protein in mechanically
wounded tobacco plants
JOURNAL Plant Cell Physiol. 41 (6), 684-691 (2000)
MEDLINE 20399434
PUBMED 10945337
REFERENCE 2 (bases 1 to 1959)
AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1997) Kojiro Hara, Nara Institute of Science and
Technology, Plant Molecular Breeding, Takayama 8916-5, Ikoma, Nara
630-0101, Japan (E-mail:k-hara@nbs.aisr-nara.ac.jp,
Tel:+81-743-72-5653 (ex.5653), Fax:+81-743-72-5659)
FEATURES
source location/Qualifiers
1..1959
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/db_xref="taxon:4097"
/tissue_type="leaf"
94..1626
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="BA095788.1"
/db_xref="GI:8096266"
/translation="MPIENFKYSPVNVKYTESEISVYDYOTTELVEHEKNGTYOMTV
KPKTVKERKDVHVPKLGVMIMGVNGKSTLTGGVLANREGISMATKDKVOANVF
GSUTASTIRVSGRBEIYAPFKSLILPVNPDVYFGMSDLSGNLADANARAYPD
IDLQKLRPYMSVPLPEIYDPFLAAGSRANNVKIGATKEEDIDDIITKIDRFKE
KKKDKVVVLTANTERTYSNVVVGINDTWNLFASVDREANISPSITVALACILENV

PFINGSPONTFVPGILDLAIKRNLTIGDDFKSGQTKKSVLVDPLVAGIKPISIVS
YNHIGNDGNMLSAPOTFRSKIEISKNVDDWMSNAILKEPHGPHVAVIKVYVY
GDSKRADEYSEIFMGKNTIVLANCEBLSLAIPIILDLVLAESTRIOLNAGE
GKFSFPHVATILSLTKAPLVPPGTVPVVALSHQRMALENILACVGLADENNMLE
YK"
1959
/note="15 a nucleotides"
ORIGIN
polyA_site
Alignment Scores:
Pred. No.: 1,078-189 Length: 1959
Score: 2475.00 Matches: 473
Percent Similarity: 96.27% Conservative: 18
Best Local Similarity: 92.75% Mismatches: 19
Query Match: 94.07% Indels: 0
Gaps: 0
DB: 8
US-10-718-952-12 (1-510) x AB009881 (1-1959)
QY 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyThrGluThrGluIle 20
DB 94 ATGTTATTTAGAACTTTAAGGTTGAGAGCCCAACGTTAAGTACACCGAAAGTGAGATT 153
QY 21 GlnSerValIlyrAsnTyThrThrThrGluLeuValHisGluAsnArgAsnGlyThrTy 40
DB 154 CACTCTGCTATGATTATCAACCACTGAGTTAGTTCATGAGAGAAAATGGGACTTAC 213
QY 41 GlnTrpIleValLysProLysSerValAsnTyThrPheLysThrAsnThrHisValPro 60
DB 214 CATGAGACTGTCACACCTTAAGACTCTCAAAATATGAGTTCAAGACTGATCTCATTTCC 273
QY 61 LysLeuGlyValMetLeuValGlyTrpGlyLysAsnAsnGlySerThrLeuThrGlyGly 80
DB 274 AATTAGGGGTTATATGCTTGTGATGGGTGAAACATATGTTCAACTTACCGGTGAT 333
QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 334 GTTATGTCTAACAGAGAAAGAAATTCATAGGCGCACAAAGATAGTCAACACCAT 393
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGly 120
DB 394 TACCTTGCTCTCTTACCTACGCTTCTACTTTCGATGGGTCTTTCAATGGAGAAAG 453
QY 121 IleTyAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 454 ATCTATGCTCAATTTAAAGCCTCTTCCATATGTCATCCAGATGAGATGATGTTTGA 513
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 514 GGATGGACATCAGCGGCATGATTTGGCAGATGCGCAAGGGCTPAAGGATTTGAT 573
QY 161 IleAspLeuGlnLysGlnLeuArgProTyTrpMetGluSerMetValProLeuProGlyIle 180
DB 574 ATTGATCTCAAAAGCAAGTTGAGGCTCTACATGAAATCTATGCTCCACTACCTGATATC 633
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 634 TATGATCCTGATTTCTATTCATGCTCTACCAAGGTCACGTGCCAACACAGTATCAAGA 693
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 694 ACCAAGAAAGAACAAATGATCAATCTTAAGATATTAGGAGATTAAAGAAAGAAC 753
QY 221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgTySerAsnLeuVal 240
DB 754 AAAGTGAACAAGGTGCTGCTGTGGACTGCTACCTGAAGATACAGTAATGTGGT 813
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
DB 814 GTTGACCTTAATGATACATAGAAATCTCTTGTCTGTGGACAGAAATGAAGCTGAA 873
QY 261 IleSerProSerThrLeuTyAlaIleAlaCysValMetGluAsnValProPheIleAsn 280

Db 874 ATATCTCTTCCACTTTGTATGCTATTGGCTGATCTTGAAATATGCTTTATTAT 933
 QY |||||SerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAnthr 300
 Db 934 GGAAGCCCAACAACACCTTTGTGCCAGGCTCATTTGATTTGGCCATTCAGAGAACACA 993
 QY 301 LeuIleGlyValAspAspPheLeuSerGlyGlnThrIleMetLeuSerValLeuValAsp 320
 Db 994 CTGATTTGGTGTGATGATCTTTAAGAGTGTCTCAACCAATGAGTCACTGCTGTGAT 1053
 QY 321 PheLeuValGlyValGlyIleLeuSerProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
 Db 1054 TTCTCTTTGGAGCTGTGATTTAAGCCACATCAATTTGATGAGTACACATTTGGTTAAC 1113
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLeuGlyIleSerLeuSer 360
 Db 1114 AATGATGGAATGAATCTGTCTGCTCCCTCAAACTTCCCTCAAGAGAGATCTCGAAAGT 1173
 QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnHisPro 380
 Db 1174 AATGTTGTTGATGACATGAGTTTCAAGCAATGCCATTTCTTATGAGCTTGAAGACCTT 1233
 QY 381 AspHisValValValIleLeuSerTyrValProTyrValGlyAspSerLeuArgAlaMetAsp 400
 Db 1234 GACCATGTTGTTGATTTAAGTATGTCATATGTGGACAGACAGCAAAAGGCAATGAT 1293
 QY 401 GlnTyrThrSerGlnIlePheMetGlyGlyValAsnThrIleValLeuHisAsnThrCys 420
 Db 1294 GAGTACACATCTGAGATTTTTCATGGAGAGAAAGAACACATTTGTTTGCACATATCTGT 1353
 QY 421 GlnAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGlnLeuSer 440
 Db 1354 GAGGATTCCTCTTTGGCTCTCCATTAATCTTGATTTGGTCTTCTTCTGTAACCTCAT 1413
 QY 441 ThrArgIleGlnPheLeuValGlnAsnGlnGlyTyrPheHisSerPheHisProValAla 460
 Db 1414 ACCCCCATTCACCTCAAACTGAGAGAGAGGTAAGTTCCATCTCTTCCACCCCGGCT 1473
 QY 461 ThrIleLeuSerTyrLeuThrIleValProLeuValProProGlyThrProValValAsn 480
 Db 1474 ACTATCCCAAGCTTACCTTCCAGAGGCTCTCTGTTGCCACACAGTACACAGTGGTGAAT 1533
 QY 481 AlaLeuSerLeuGlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAla 500
 Db 1534 GCACCTCTCAAGCAGAGGCAATGCTTGAGAACATATGAGGCTTGTGTGACTTGA 1593
 QY 501 ProGlnAsnAsnMetIleLeuGlnTyrIleCys 510
 Db 1594 CCAGAGAACACATGATTTCTGGAATTCATA 1623
 RESULT 6
 AB032073 1931 bp mRNA linear PLN 08-SEP-1999
 LOCUS Nicotiana paniculata NpINP51 mRNA for myo-inositol-1-phosphate
 DEFINITION synthase, complete cds.
 ACCESSION AB032073.1 GI:5834499
 VERSION myo-inositol-1-phosphate synthase.
 KEYWORDS Nicotiana paniculata
 SOURCE Nicotiana paniculata
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; Lamiales; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (bases 1 to 1931)
 AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
 JOURNAL Myo-inositol-1-phosphate synthase
 REFERENCE Published Only in Database (1999)
 AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
 JOURNAL Direct Submission
 TITLE Submitted (01-SEP-1999) Akiko Hashimoto, Japan Tobacco Inc., Plant
 Breeding and Genetics Research Laboratory; 700 Higashibara, Iwata,
 Toyoda-cho, Shizuoka 418-0802, Japan

(E-mail: Akiko.Hashimoto@pbgrl.jti.co.jp, Tel: +81-538-32-7116,
 Fax: +81-538-33-6046)
 FEATURES
 source Location/Qualifiers
 1..1931
 /organism="Nicotiana paniculata"
 /mol_type="mRNA"
 /db_xref="taxon:62141"
 1..1931
 /gene="NpINP51"
 /gene="NpINP51"
 92..1624
 /gene="NpINP51"
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase"
 /protein_id="BA84084.1"
 /db_xref="GI:5834500"
 /translation="MEIENFKVESPNVRYTSEHSIVDYQTTLEVHDEKNGTYQMTV
 KPTKYKFEKDVHPKGLVIMVGGNNGSTLTGVLANRGIISMAKDKQOANYF
 GSLTQASTLRVGSFNGSEIVAPFKSLPMVNPDDVVFEGMDISDMNLADMAARAFD
 IDLQKSLPVMESVPLPGIYDDPDLIANGSRRANVYKGTKEQIDQIKIDREEK
 KKQDVVVLTMTANTRYSNVYVGNMTMENIPASVDNREIISSTLYAICLENV
 PFINSPQNTFVPLIDLIKKNTLIGDDDFKSGOTRKSVLVDLVAGIKPTSLV
 YNHLGNNDGMNLAPQTFRSKEISKNVDDMVSSNALYEGEHPDHVVVTKYVYV
 GDSKRAMDEYSEIFMGKNTIIVLHNTCEDSLAIPILIDLVLAEISTRIDLKAGE
 GKFSHFHVATILSYLTKAPLVPGTPTVNALSKRAMLENILRACVGLAPENNMLE
 YK"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,656-189 Length: 1931
 Score: 2470.00 Matches: 471
 Percent Similarity: 96.47% Conservative: 21
 Best Local Similarity: 92.35% Mismatches: 18
 Query Match: 93.88% Indels: 0
 DB: 8 Gaps: 0
 US-10-718-952-12 (1-510) x AB032073 (1-1931)

QY 1 MetPheIleGlnAsnPheLeuValGlnSerProAsnValIleTyrThrGlnThrGlnIle 20
 Db 92 ATGTTATTGAGAAATTTTAAGGTTAGAGCCCAACGTTAATACACCGAAATGTAAT 151
 QY 21 GlnSerValTyrAsnTyrGlnThrThrGlnLeuValHisGlnAsnArgAsnGlyThrTyr 40
 Db 152 CACTCTGTCTATGATTAACAACCTGAGTTAGTTCATGATGAGAAAATGGACATAT 211
 QY 41 GlnTPleValIleAspProIleSerValAsnTyrGlnPheLeuThrAsnThrHisValPro 60
 Db 212 CAATGACCGTCAACGCTTACACTGCAAAATATGATGTTCAACACTGATGTTGATCCC 271
 QY 61 LysLeuGlyValMetLeuValGlyTyrPqlGlyAsnAsnGlySerThrLeuThrGlyGly 80
 Db 272 AAATTAGGGGTTATCTGTTGATGGGTGGAACCAATGTTCAACCTTGACCGGTGCT 331
 QY 81 ValIleAlaAsnArgGlnAspIleSerTyrAlaThrIleAspValIleGlnGlnAlaAsn 100
 Db 332 GTTATTGCTACAGAGAAATTTCAATGGCCACCAAGATAGAGTGCACAGCCCAAT 391
 QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlnIle 120
 Db 392 TACTTTGGGCTCTTACTACAGGCTTCTACTATGCAAGTTGGGCTTTCAATGAGAGAG 451
 QY 121 IleTyrAlaProPheLeuSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 Db 452 ACTATGCTCCATTTTAAAGCTCTCTTCAATGTCATCCAGATGAGAGCTACTGTTTGA 511
 QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 Db 512 GGATGGACATCAGGACATGAAATTTAGCAGATGCCAGCCAGGCTTAAGGTAATTTGAT 571
 QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlnSerMetValProLeuProGlyIle 180
 Db 572 ATGATCTACAAAGCAGTGGGCTCATATGAGAACTATAGTGTCCCACTCCCTGGTATC 631

QY 181 TyrAspProAspPheIleAlaIaIaengIngluArgAlaAsnAsnValIleYsgly 200
DB 632 TATGACCTGATTTATGCTGCTCTAACCAGGCTCAGCGCAACAGCTGATCAAGA 691
QY 201 ThrIysGIngluGlnValGlnGlnIleIleYsAspIleYsAlaPheIysGluAlaThr 220
DB 692 ACCAAGAAAGAAACAATTGATCAATTAAGATTAATAGGAGCTTTAGCAAAAAGAC 751
QY 221 LysValAspLysValValValLeuThrPheIleAsnThrGlnArgTyrSerAsnLeuVal 240
DB 752 AAAGGAGCAAGGAGTGTATGTGAGCTGCTAACACTGAAAGATACAGTATGTGGTT 811
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
DB 812 GTTGGACCTTAATGACATATGAGAAACCTTTGCTTCTGTGACAGAAATGAACTGAA 871
QY 261 IleserProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 872 ATATCTCCTTCCACTTGTATGCTATGTGCTGCAATCTTGAAATGTGCTTTATTAAT 931
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 932 GGAAAGCCCCAGAAACCTTTGTCCAGGCTCATTTGATTTGGCCATCAAGAAACACA 991
QY 301 LeuIleGlyGlyAspAspPheIysSerGlyGlnThrIysMetIysSerValLeuValAsp 320
DB 992 TTGATTTGGTGGTGAAGCTTTAAAGAGTGGTCAACAAATGAATCAAGTGTGTTGAT 1051
QY 321 PheLeuValGlyAlaGlyIleYsProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 1052 TTCTCTGTGGAGCTGATTAATGAACCAATCAATTTGATGACATCAACCATTTGGGTAA 1111
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIysGlnIleSerIysSer 360
DB 1112 AATGATGGAGTAATGTCTGTCCCTCAACTTTCCGCTCAAGAGAGATCTGAAAGT 1171
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1172 AATGTTGTTGATGACATGCTTTCAAGCAATGCCATCTTTAGAGCTGAGAGACCTT 1231
QY 381 AspHisValValValIleYsTyrValProTyrValGlyAspSerIysArgAlaMetAsp 400
DB 1232 GACCATGTTGTTGATTAAGTATGAGCCATATGTGGAGACAGCAAGGCGCAATGAT 1291
QY 401 GlyTyrThrSerGlnIlePheMetGlyGlyYsAsnThrIleValLeuHisAsnThrCys 420
DB 1292 GAGTACACATCTGAGATTTTCAATGGGGGAAAGAACACCATTTGTTGCAATATCTGT 1351
QY 421 GluAspSerLeuLeuAlaIaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1352 GAGGATTCACCTTTAGCTGCTCCAAATTATATGATTTGGTCTCTTGTCTGAATCACT 1411
QY 441 ThrArgIleGlnPheIysAlaGluAsnGluGlyYsPheHisSerPheHisProValAla 460
DB 1412 ACCCCCATTCACCTCAAGCTCAAGAGAGGAGGTAAGTTCCACTCTCCACCCCTGGCT 1471
QY 461 ThrIleLeuSerTyrLeuThrIysAlaProLeuValProProGlyThrProValIleAsn 480
DB 1472 ACTATCTCAGCTACCTTAACCAAGGCTCTCTGTGTAACCAAGGTAACCAATGTGGAT 1531
QY 481 AlaLeuSerIysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1532 GCACTCTCAACAGACAGGCAATGCTTGAGAACATATTAAGGGCTTTGTGTGACTTGA 1591
QY 501 ProGluAsnAsnMetIleLeuGluTyrIys 510
DB 1592 CCAGAGAACCAACATGATCTGGAATACAA 1621
RESULT 7
E27176 1950 bp DNA linear PAT 18-JUN-2001
LOCUS E27176
DEFINITION Novel INPS gene derived from nicotiana.
ACCESSION E27176

VERSION E27176.1 GI:13026394
KEYWORDS JP 1999187879-A/1.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE
AUTHORS Shigenori, Y. and Toshiyuki, K.
TITLE Novel INPS gene derived from nicotiana
JOURNAL Patent: JP 1999187879-A 1 13-JUL-1999;
JAPAN TOBACCO INC
COMMENT
OS Nicotiana paniculata
PN JP 1999187879-A/1
PD 13-JUL-1999
PF 26-DEC-1997 JP 1997359773
PR SHIGENORI YAMADA, TOSHIYUKI KOMORI
PI C12N15/09//C12N15/10, (C12N15/09, C12N15/91), C12N15/00, C12N15/00,
PC (C12N15/00, C12N15/91)
CC Strandedness: Double;
CC Topology: Linear;
CC Key Location/Qualifiers
FH Key 92..1703.
FT CDS Location/Qualifiers
1..1950
/organism="Nicotiana paniculata"
/mol_type="genomic DNA"
/db_xref="taxon:62141"
ORIGIN
Alignment Scores:
Pred. No.: 2,696-189 Length: 1950
Score: 2470.00 Matches: 471
Percent Similarity: 96.47% Conservative: 21
Best Local Similarity: 92.35% Mismatches: 18
Query Match: 93.88% Indels: 0
DB: 6 Gaps: 0
US-10-718-952-12 (1-510) x E27176 (1-1950)
QY 1 MetPheIleGluAsnPheIysValGluSerProAsnValIysTyrThrGluThrGluIle 20
DB 92 ATGTTATTTAGAAATTTTAAAGTTGAGAGCCCAACGTTAAGTACACCGAAAGTGAATT 151
QY 21 GlnSerValIYrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 152 CACTCTGCTATGATTAATCAACCACTGAGTTAGTTCATGATGAGAAAAATGGGACATAT 211
QY 41 GlnTPrIleValIlyPProIysSerValAsnTyrGlnPheIysThrAsnThrHisValPro 60
DB 212 CAATGACCGCTCAAGCCCTTAAGACTGTCAAAATATAGATTCAAGACTGATGATGTTCC 271
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 272 AAATTAGGGTTATATCTGTTGATGGAGGGGTGAACATGGTTCAACCTTGACCGGTGCT 331
QY 81 ValIleAlaAsnArgGluAspIleSerTyrAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 332 GTTATTTGCTAACAGAGAAAGAAATTTCAATGGGCCAACCAAGATTAAGTCCAAAGCCCAT 391
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
DB 392 TACTTTGGCTCTCTTACTCAAGCTTTCTACTTTCGATGGGCTCTTTCAATGGAGAAAG 451
QY 121 IleTyrAlaProPheIysSerIleuLeuProMetValAsnProAspAspIleValPheGly 140
DB 452 ATCTATGCTCATTTAAAGCCTCTCTTCCAAATGCTCAATCCAGATGACGATGTGTTGGA 511
QY 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 512 GGATGGACATCAAGCAGCATGATTAATTTAGCAAGATGCCATGGCCAGGGCTTAAGTATTGAT 571

```

OY 11eapleuglnlysglnleuargProGlyrMetGluSerMetValProleuProGlyIle 180
DB 572 ATTGATCTCAAAAGAGAGTTGAGCCCTTACATGATCTATGTTCCCATCTCCCTGATATC 631
OY 181 TTAAspProaspPheIleAlaIaenglngluGluArgAlaasnValIlelysgly 200
DB 632 TATGACCTGATTTCAATGCTGCTAACCAAGGCTCAGTCCCAACACGTGATCAAGGA 691
OY 201 TTTlysglngluGlnValglnGlnIleIlelyAspIlelyAspIlelyAspIlelyAspIle 220
DB 692 ACCAAGAAAGAACAAATGATCAAAATCAATTAAGGATTAAGGAGTTTAAGGAAAAAAC 751
OY 221 LysValAspLysValValIleuTPPrAlaAsnThrGluArgThrSerAsnLeuVal 240
DB 752 AAAGTGGACAAAGGTGATGATTTGAGCTGCTAACACGAAAGATACGATATGTGTT 811
OY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
DB 812 GTTGGACTTAATGACACTATGGAAAACCTTTGCTTGTGTGACAGAAATGAGCTGA 871
OY 261 TTAAspProaspPheIleAlaIaenglngluGluArgAlaasnValIlelysgly 280
DB 872 AATATCTCTTCCACTTTGATGCTATGCTGCTGATCTTGAAGAAATGCTTTTATTAT 931
OY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 932 GGAAGCCCCAGAACACCTTTGTCAGAGCTCATGATTTGGCCATCAAGAAAGACACA 991
OY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValIleuValAsp 320
DB 992 TTGATTTGGTGGTATGATCTTTAAGAGTGTCAACCAAAATGAAAGTCAGTGGTGTAT 1051
OY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerThrAsnHisIleuGlyAsn 340
DB 1052 TTCTCTTTGGTGGCTGCTGATTAAGCCAAATGATGATGATGATGATGATGATGATGAT 1111
OY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
DB 1112 AATGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1171
OY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrgluProGlyIleHisPro 380
DB 1172 AATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
OY 381 AspHisValIleValIleLysTyrgluProTyrgluValGlyAspSerLysArgAlaMetAsp 400
DB 1232 GACCAATGTTGTTGATTAAGTATGTCATATGTCGAGACAGCAAGAGGGCAATGGAT 1291
OY 401 GlnTyrgluThrSerGlnIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
DB 1292 GAGTACACATCTGAGATTTTCATGGGGGGAAGAAACACCAATGTTTTCACAAATCTTGT 1351
OY 421 GlnAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1352 GAGGATTCATCTTTACTGCTCCAAATTAATGATTTGCTCTTCTGCAACCTCAAT 1411
OY 441 ThrArgIleGlnPheLysAlaGluAsnGlyGlyLysPheHisSerPheHisProValAla 460
DB 1412 ACCCGCATTCACCTCAAGCTGAAAGAGAGGATTAAGTTCACCTTCCACCCCGGAGCT 1471
OY 461 ThrIleLeuSerTyrgluThrLysAlaProLeuValProProGlyThrProValIleAsn 480
DB 1472 ACTATCTCAGCTTACCTTACCAAGGCTCTCTGTGATCAACCAAGTACACCAAGTGGAGAT 1531
OY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1532 GCACTCTCAAGAGAGGGCAATGCTTGAAGACATTTGAGGGCTTGTGTGCACTTGA 1591
OY 501 ProGluAsnAsnMetIleLeuGlyTyrgly 510
DB 1592 CCAGAGAACACATGATCTCGAATACAAA 1621

```

```

BT013505
LOCUS      BT013505      1959 bp      mRNA      linear      PLN 11-MAY-2004
DEFINITION      Lycopersicon esculentum clone 132193F, mRNA sequence.
ACCESSION      BT013505
VERSION      BT013505.1 GI:47104920
KEYWORDS      FLI CDNA.
SOURCE      Lycopersicon esculentum (tomato)
ORGANISM      Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE      1 (bases 1 to 1959)
AUTHORS      Kirkness, E.F., Mang, W. and Vazelle, A.
TITLE      Direct Submission
JOURNAL      Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
              Medical Center Drive, Rockville, MD 20850, USA
FEATURES
     source          1..1959
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4081"
                     /clone="132193F"
                     /tissue_type="fruit"
                     /note="TZBFS1"
ORIGIN
Alignment Scores:
Pred. No.:      3,92e-189      Length:      1959
Score:          2468.00      Matches:      472
Percent Similarity: 96.27%      Conservative: 19
Best Local Similarity: 92.55%      Mismatches:  19
Query Match:    92.80%      Gaps:        0
DB:
US-10-718-952-12 (1-510) x BT013505 (1-1959)
OY 1 MetPheIleGluAsnPheLysValGluSerProAsnValIleTyrgluThrGluThrGluIle 20
DB 77 ATGTTATATGAAATTTTAAGGTGGAAAGCCCAATGTGAATATATGAGAAATGAAATTT 136
OY 21 GlnSerValTyrgluAsnTyrgluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyrg 40
DB 137 CATTCGTGTATGATTAATGAACCAAGAGCTGTTTCATGAAGAGAAATGAACTTAT 196
OY 41 GlnTPPleValIleLysProLysSerValAsnTyrglnPheLysThrAsnThrHisValPro 60
DB 197 CAATGATGTTTAACTTAACTGTAATATGAATTTTAAACTGATACCATGAGGCCA 256
OY 61 LysLeuGlyValMetLeuValGlyTyrglyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 257 AAATTGGGGTTATCTTGTGATGAGGAGGAAACAAATGTTCAACATTCGAGGT 316
OY 81 ValIleAlaAsnArgGluAspIleSerTPPrAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 317 GTTATTCGCAATCGAAGAGATTTTCATGCGCAAGAAAGAAAAGTGCACAAACCAT 376
OY 101 TyrgluGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGln 120
DB 377 TATTTTGGGTCTTAACTCAGGCAATCAACATTCAGATTGAGGTCTTTCATATGGCGAAGAG 436
OY 121 IleTyrgluProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 437 ATCTATGACCCCTTAAAGCTCTCTCCCAATGTCACCCAGAGAGATGATATTTTGA 496
OY 141 GlyTPPrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 497 GGAATGCAATTAAGCAATCATGAATTTGCGAGATGCTATATGTAAGGCTTAAGTTTTCGA 556
OY 161 11eapleuglnlysglnleuargProGlyrMetGluSerMetValProleuProGlyIle 180
DB 557 GTTATCTGCAAAACACAGCTCAGGCCCTTACATGATCATCATGATCCCTCTCTGTATC 616
OY 181 TyrgluProaspPheIleAlaIaenglngluGluArgAlaasnValIlelysgly 200

```



```

Db      617 TATGAGCCCTGACTTCATTCGGCTTAACCAAGAGCAGCGCCAAACCTGATCAAGAA 676
Qy      201 ThrValGlnGluValGlnGlnIleIleValAspIleLeuValAlaPheValGluAlaThr 220
Db      677 ACCAAGAAAGAAACAACTTCAACAAATTTGTTAAAGATATTAGGAGATTCAAGAGAAAGAAC 736
Qy      221 LysValAspValValValLeuThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      737 AAGGTAGACAAAGATAGTGGTTCTATGACTGCCAACACCGAAAGGTACGTAATGGTT 796
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
Db      797 GTTGGCTTAATGACACCAATGAAAACTTTTAGCTGCTGATGATGAATAGAGCTGAA 856
Qy      261 ILeuSerProGlnThrLeuThrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      857 AATATCTCTTCTTACATTCATGATGCTATGCTGTATATATGAAAAGTGGCTTTCATCAAC 916
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      917 GGAAGCCCTCAAAACACTTTTGTCCAGGCTTATGATTTGGCCATTAAGAAAGAACACT 976
Qy      301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      977 TTAATTTGGTGTGATGACTTAAAGTGTCAACCAACAGATGAAGTCCGTGTGGTTGAT 1036
Qy      321 PheLeuValAlaValAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db      1037 TTCTCTTTGGAGCCGGATTTAAAGCCACCTCAATATGTAAGCTACATTAAGTGGTAAAC 1096
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db      1097 AATGATGGAATGATCTTCTGCTGCTCAACCTTCGGCTTAAGAGATCTCAAAAGAT 1156
Qy      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db      1157 AATGTTGTGATGACATGCTGTCTAGCAACGCCATCTCTATGAATCTGGCAGACACCT 1216
Qy      381 AspHisValValValIleLysTyrValProTyrValAlaAspSerLysArgAlaMetAsp 400
Db      1217 GACCATGTTGTGTATCAAGATATGTTCCATATGTTGGAGACAGCAAGGGCAATGAT 1276
Qy      401 GlnTyrThrSerGlnIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
Db      1277 GAGTACATGTCAGAGATTTTTCATGGCGGAAAGACACTATAGTATGCAACAACCTGT 1336
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1337 GAGGACTCTCTTTGGACAGCTCCAATTATCTTGGAATTTGTCCTCTCGTCAACCTCAC 1396
Qy      441 ThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      1397 ACTCGCATTCAGCTCAAAAGCTGAAGGGAGGGAGAACTTCATCTCTTCAACCTGTGGCT 1456
Qy      461 ThrIleLeuSerTyrIleuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db      1457 ACTATATCTAGCTACTCAACCAAGGCTCCCTGGTATCCACCAAGTACTCCAGTGGGAT 1516
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1517 GCCCTTTCAAGCAGAGGGCAATGCTTGAGAAATATATAGAGAGCTTGTGGTTGGCA 1576
Qy      501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db      1577 CCAGAGAACACATGATTTTGGATACAA 1606

```

RESULT 9 AF284065 1845 bp mRNAs linear PIN 30-JUL-2003
 AF284065 LOCUS Sesamum indicum myo-inositol 1-phosphate synthase mRNA, complete
 DEFINITION Cds.
 ACCESSION AF284065

```

VERSION AF284065.1 GI:3958815
KEYWORDS Sesamum indicum (sesame)
SOURCE Sesamum indicum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asceridales; Lamiales; Pedaliciales; Sesamum.
REFERENCE 1 (bases 1 to 1845)
AUTHORS Chun,J.A., Jin,U.H., Lee,J.W., Yi,Y.B., Hyung,N.I., Kang,M.H.,
Pyee,J.H., Suh,M.C., Kang,C.W., Seo,H.Y., Lee,S.W., and Chung,C.H.
Isolation and characterization of a myo-inositol 1-phosphate
synthase cDNA from developing sesame (Sesamum indicum L.) seeds:
functional and differential expression, and salt-induced
transcription during germination
JOURNAL Planta 216 (5), 874-880 (2003)
MEDLINE 22511750
PUBMED 12624775
REFERENCE 2 (bases 1 to 1845)
AUTHORS Jin,U.-H., and Chung,C.-H.
Direct Submission
JOURNAL Submitted (01-JUL-2000) Division of Biotechnology, Faculty of Life
Science and Resources, Dong-A University, Ha-Dan-Dong, Sa-Gu-Gu.,
Pusan City, Pusan 604-714, South Korea
FEATURES
source
1..1845
/organism="Sesamum indicum"
/mol_type="mRNA"
/db_xref="taxon:4182"
/tissue_type="seeds"
/note="authority: Sesamum indicum L."
51..1583
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AA01148.1"
/db_xref="GI:3958816"
/translation="MPFESFVSESFNKYTEGEIHSVYNYETTELVAHSRNGTYOMIV
KEKTVKEFKTDHVPKLGKMLVCGKNGSTLTGCTIARREGISNATKDYQANVF
GSLVQASSIRVGSFGEETIAPFKSLIPVAPDVVGWDISNMNLADAPGRKVLQ
IDLOKOLRPFVMEHWPLPGIYDPDFIANQSSRANNYIKTKKQVOQIIKDMDFKE
ONKDKVVLVLTATERSNVVGLINPTAESLMAVSRENEMISPSLTALACVENV
PINGSQNTFVPGILIDAIORNSLIGDDSKSQTKSVLVDPLVGAGIKPTSIYS
YNHGNDGMLSPQPFERSKETSKSNVVDVMVASNGILVRSRPHIIVIKVPPVY
GDSKRADEYTSIFMGSKSTIVLANTCEBSLAPILIDVLAEISTRQLQAREG
GKFSFHEVATILISLYTAPLVPGTIPVVALSKORAMLENILRACVGLAENNNILR
YK"
ORIGIN
Alignment Scores:
Pred. No.: 7,06e-188 Length: 1845
Score: 2452.00 Matches: 470
Percent Similarity: 96.08% Conservative: 20
Best Local Similarity: 92.16% Mismatches: 20
Query Match: 93.20% Indels: 0
DB: Gaps: 0
US-10-718-952-12 (1-510) x AF284065 (1-1845)
Qy      1 MetPheIleGluAsnPheLeuValGluSerProAsnValLysTyrThrGluThrGluIle 20
Db      51 ATGTTCAATCGAGAGCTTCAAGGTTGAGAGCCCAATGTCAAGTACACCGAGGTGAGATT 110
Qy      21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db      111 CACTGTGCTCAACTACAGACTACAGAGCTTGTTCATGAGAGAGAAAGGCACTTAC 170
Qy      41 GlnTPIleValLysProLysSerValAsnTyrGlnPheLysTyrAsnThrHisValPro 60
Db      171 CAATGATTTGCAAGCCCAAGACTGCTCAATAAGAGTTCAAAACCTGATACATGATCCCC 230
Qy      61 LysLeuGlyValMetLeuValGlyTTPGlyGlyLysAsnArgLysSerThrLeuThrGlyGly 80
Db      231 AATTTAGGGGTGATGCTTGTGGATGGGAGGAAACAATGATTCATCACTCACTGGCGGT 290

```



```

QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrIleAspIleGlnGlnAlaAsn 100
DB 291 GTCATTGCCAAACAGGAGGATTTTCATGGGCAACAAAGATAGGTGCACGGCCAAAT 350
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGln 120
DB 351 TATTTTGGGCTCAGTGCACAGGATCTCATCATGAGTTGGTCTTCAATGAGAAAGAG 410
QY 121 IleTyrAlaProPheIleSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 411 ATCTATAGCCCTTCAAAAGCTTCTTCCATGGTCAACCCAGATACGATAGGTTGGG 470
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
DB 471 GGATGGGATCATCGCAATATGAAATTTGGCTGATGCCATGGCAGGGCCAAAGGTTAGAT 530
QY 161 IleAspLeuGlnIleGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 531 ATTTGATCTCCAGAGCAGCTCAGGCCCTCATGAGACACATGGTCCACTCCCTGGAAATC 590
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluGlnAlaAsnValIleIleGly 200
DB 591 TACGATCTGATTTTCAAGCCGCAATCGAGATCAAGTCCAAACGATGATTAAGA 650
QY 201 ThrIleGlnGlnValGlnGlnIleIleIleAspIleIleValPheIleGlnAlaThr 220
DB 651 ACCAAGAAAGAACAGATTCAACATCATCAAGACATGAGGACTTTAAGAGCAAAAT 710
QY 221 IysValAspIleValValIleuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 711 AAGGTGCACAGGCTGATGCTGATGTCAGACAGCAACAGAAAGATACGATATGTTGTC 770
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaValAspArgAsnGluAlaGlu 260
DB 771 GTTGGCTTAAAGATACACCGCAAGCTGATGCTCGTGGAGAGATGAGCGGAG 830
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 831 ATATCTCTTCAACCTTGATGCAATGCTGATGTTGTTGAAATGTTCCGTCATTAAT 890
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 891 GGCAGCCCAAAATATCTTTGTTCAGGGTATGATTTGGCGATTCAAGAGAAACAT 950
QY 301 LeuIleGlyGlyAspAspPheIleSerGlyGlnThrIleMetIleSerValLeuValAsp 320
DB 951 TTGATTTGGGAGATGATTTCAAGCGGCTCAAAACCAAGATGAATCACTGCTAGTTGAT 1010
QY 321 PheLeuValGlyAlaGlyIleIleLysProThrSerIleValSerTyrAsnHisIleGluIleAsn 340
DB 1011 TTCTCTTGTGGAGCTGGATTCAGGCCCAAGCTCATTTGATGATATATCATCTGGGAAAC 1070
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGlyIleSerIleSer 360
DB 1071 AACGACGGGATATCTATCAACACCCCAACCTTCCGTCCTCAAGAGATTTCTAAAC 1130
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1131 AATGTGTAGACAGCATGCTTCCAGCAACGACATCTTATGAGCTGTGTAACACCCC 1190
QY 381 AspHisValValIleIleIleTyrValProTyrValGlyAspSerIleArgAlaMetAsp 400
DB 1191 GACCAATATTTGTCTCAAGATGACCGTATGTGGGAGACAGCAAGACCAATGAGC 1250
QY 401 GlyTyrThrSerGlnIlePheMetGlyGlyIleAsnThrIleValLeuHisIleAsnThrCys 420
DB 1251 GAGTACACTTCCGAGATATTCATGGAGGAAAGATACCATAGTTCTGCACACACGTGT 1310
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1311 GAGGACTCTCTTCTGCTGCTCCCATCATCTGAGATTTGTGCTCTCTAGCTAACTACAGC 1370
QY 441 ThrArgIleGlnPheIleValGluAsnGluGlyIlePheHisSerPheHisProValAla 460

```

```

DB 1371 ACTCGATTCACACTCAAGCCAGGAGGAGGCAAAATTCATTCCTGCCACCCGGGCA 1430
QY 461 ThrIleLeuSerTyrLeuThrIleValProLeuValProProGlyThrProValValAsn 480
DB 1431 ACCATCTCTAGCTACCTAACAAGCCCTTGTGATCTCCAGGACACCGGTCGTCAAC 1490
QY 481 AlaLeuSerIleGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1491 GCACTGTCTAAGCAGGCGCCAGCTCGAAGACATCTTGAGGCTTGTGTGATTTGGCT 1550
QY 501 ProGluAsnAsnMetIleLeuGlyTyrIle 510
DB 1551 CCAGAAACAAATATGATTTTGGATATCAAG 1580

RESULT 10
BT013759 1954 bp mRNA linear PLN 11-MAY-2004
LOCUS Lycopersicon esculentum clone 132642F, mRNA sequence.
DEFINITION BT013759
ACCESSION BT013759.1 GI:47105174
VERSION FLI_CDNA.
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum
REFERENCE 1 (bases 1 to 1954)
AUTHORS Kitzness, E.F., Wang, W. and Vazeille, A.
TITLES Direct Submission
JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA

FEATURES
source 1..1954
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132642F"
/cisue_type="fruit"
/note="TZBX41"

ORIGIN
Alignment Scores:
Pred. No.: 1,226-186 Length: 1954
Score: 2437.00 Matches: 464
Percent Similarity: 95.49% Conservative: 23
Best Local Similarity: 90.98% Mismatches: 23
Query Match: 92.63% Indels: 0
DB: 8 Gaps: 0

US-10-718-952-12 (1-510) x BT013759 (1-1954)
QY 1 MetPheIleGluAsnPheIleValGluSerProAsnValIleTyrThrGluThrGluIle 20
DB 107 ATGTTATTTAGACACTTTTAAGGTTAGAGCCCAATGTTAAGTACACAGAGTGAAGATT 166
QY 21 GlnSerValIleTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 167 CACTCTGTGTATGATTTAACAACACTGATGTTGTTCAATGAGAAATAAATGGACATTT 226
QY 41 GlnTrpIleValIlePheProLysSerValAsnTyrGlnPheIleThrAsnThrHisValPro 60
DB 227 CAATGAGCTGTAAACCCAAAACGTGCAAAATATGAATTCAAAACGATGTTCAATGTTCCA 286
QY 61 IysLeuGlyValIleMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 287 AATTTAGAGATTAATCTTGTGGATGGGAGTGAACAATGTTTCAGCTTACCCGAGAGT 346
QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrIleAspIleGlnGlnAlaAsn 100
DB 347 GTTATTTGCTAATAGAGAAATTTCAATGGGCTACCAAGATTAAGGTGCAGCAAGCAAT 406
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValIleGlySerPheGlnGlyGln 120

```

Db 407 TACTTTGGTTCTCTTACACAGGCCCTCTACTATTTCAGATGGGTCTTTCAATGAGAGAG 466
Qy 121 IIEYRALPProPhySerLeuPProMetValAsnProAspAlaIleValIleGly 140
Db 467 ATCTATGCTCCCTTTTAAGACACTCTTCCATGCGTCAATCCAGATGATGATTTTGGTA 526
Qy 141 GIYTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValIleAsp 160
Db 527 GGATGGGACATCAGCAACATGAAATTTGGCAGATGCATGCAGCAAGCTAAAGTTTGGAC 586
Qy 161 IIEAspLeuGlnIleGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
Db 587 AATGATCTTCAAAATGCAATGAGCGCGTACAGATGATCATGCTCCACTGCTGCTATC 646
Qy 181 TYRAspProAspPheIleAlaIleAsnGlnIleIleIleYAspIleYAspAlaPheYAspGlnAlaThr 200
Db 647 TATGACCTGATTTCTTCTGCAACCAAGGCTACGCTACCAATCTCATCAAGAAG 706
Qy 201 ThrYAspGlnGlnIleValGlnGlnIleIleYAspIleYAspAlaPheYAspGlnAlaThr 220
Db 707 ACCAAGAAGAACAAATGATGCAATTTATTAAGATATTAGGAGATTTAAGAGAACAAAC 766
Qy 221 LysValAspIleValIleValIleUtrPThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 767 AAGTGACAGCGGTGGTGGTCTGCTGAGCTGCCAACACAGAAAGATACGCGTGGCT 826
Qy 241 VALIleGlnLeuAsnAspThrMetGluAsnLeuAlaIleValAspArgAsnGlnAlaIle 260
Db 827 GTTGGCTTTAATGATACCAAGAAACCTTCTTGGCTTCTGATGATGAATAGAGCTGAA 886
Qy 261 IIESerProSerThrLeuYAlaIleAlaCysValMetGluAsnValProIleIleAsn 280
Db 887 AATATGCTCTTCAACCTTGTATGCTATGCTGATTTCTGAAATAGTACCTTTCATCAT 946
Qy 281 GlycerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 947 GGAAGGCCCAAAATGATCTTGTCCAGGCGTCATGATTTGGCCATCAAGAGAACACT 1006
Qy 301 LeuIleGlnIleYAspAspPheIleYAspIleYAsnIleUtrPThrMetIleYAspValIleValAsp 320
Db 1007 TTGATTTGGTGGTGAATCTTTTAAGAGTGTCAAAACCAATGAATGATCAAGTGGTGGT 1066
Qy 321 PheLeuValGlnIleGlnIleYAspProThrSerIleValSerTyrAsnIleLeuGlnAsn 340
Db 1067 TCCCTTTGTTGAGCGTGTATTAAGCCCAATCATCATGATGATCAACCATTTGGGTAC 1126
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnIlePheArgSerIleGlnIleSerIleSer 360
Db 1127 AATGATGGAATGATCTGCTGCCCTCCAACTTCCGTTCAAGAGAGATCTCGAAGAAC 1186
Qy 361 AsnValIleValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnIleIlePro 380
Db 1187 AACGTTTGTGATGACATGCTTCTTACTAGCCATCTTTATGAGCTGGAAGCATCTCT 1246
Qy 381 AspHleValIleValIleYAspTyrValProTyrValGlnYAspSerIleYAspAlaMetAsp 400
Db 1247 GACCATGTTGTTGTATTAAGTACCATATGTCGAGATAGCAAGAGGCGCAATGAT 1306
Qy 401 GIUTYrThrSerGlnIlePheMetGlyIleYAsnThrIleValIleUtrIleAsnThrCys 420
Db 1307 GAGTACACATCAGAGATTTTTCATGGGTGAGAGAACACCATTTGATTCACATACCTGT 1366
Qy 421 GlnAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleLeuAlaIleLeuSer 440
Db 1367 GAGGATTTCTTTTGGCTGCTCCAAATTAATTTGATTTAGTCTCTGCTGATCACTCAT 1426
Qy 441 ThrArgIleGlnPheYAspIleGlnIleGlnIleGlnIleYAspPheIleSerPheIleProValAla 460
Db 1427 ACCCGATTCAGTCAAAAGTGAAGAGAGGGGATTTCCATCTTTCATCCCGTGGG 1486
Qy 461 ThrIleLeuSerTyrLeuThrIleYAspIleProLeuValProProGlyThrProValIleAsn 480

Db 1487 AACATCTCACTTATCTTACCAAGGCTCTTGGTACCAACAGGTACACCATGCTGAT 1546
Qy 481 AlaleuSerIleYAspIleValIleMetLeuGlnIleValIleMetAlaIleValIleGlyLeuAla 500
Db 1547 GCCCTTTCAAGAGAGGCGCAATGCTGATGAACATATTAGAGGCTTGTGTTGACTGCA 1606
Qy 501 ProGluAsnAsnMetIleLeuGlnIleTyrIleYAsp 510
Db 1607 CCAGAGAACCAACATGATCTTGAATACAAA 1636
RESULT 11
AF120146
LOCUS AF120146 1538 bp mRNA linear PLN 23-APR-1999
DEFINITION Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
synthase (MPS) mRNA, complete cds.
ACCESSION AF120146
VERSION AF120146.1 GI:4589061
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 1538) Hussain, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
CDNA clone for myo-inositol 1-phosphate synthase from wheat
Unpublished
2 (bases 1 to 1538) Hussain, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
Direct Submission
Submitted (14-JAN-1999) National Research Council Canada, Plant
Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N
0W9, Canada
FEATURES
Location/Qualifiers
source 1..1538
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Fielder"
/db_xref="taxon:4565"
gene 1..1538
/gene="MIPs"
1..1533
/EC_number="5.5.1.4"
/function="converts glucose 6-phosphate into myo-inositol
1-phosphate"
/note="cytosolic enzyme"
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAD26330.1"
/db_xref="GI:4589062"
/translation="MFIESFYVESPNVXYTENIHSVYDEETFEVHENRNGTYQWV
KPKTVKDFKTRFVPLKGLVLMVGKMGNGSTLTAGVYANKEGISMAIKDKVQOANYF
GSLIQASSIRGYSNGBEIYAPEKSLIPMNVPDIVRGMDISMNILADANARAYLD
IDLOKQLPVMEHWPPUPGIYDPPIAANGSRANSYIKGTKEQVQDQITADMEPKE
KNKDKLVLTANTATERSDVVGALNDMENTLASVEKNSEISPSITVALACVLEGI
PFINGSPONTFVPLGILBLASKNCTIGDDDFKSGQTKKSVLVDFLVGAGIKPTSIYS
YNHIGNNDGMNLSAPQTFRSKEISKSNNVDVMSNGILFEFGHPDHVVYIKVPPV
GDSKRAMEYVSEIFMGSKNTIYVNTNCEBLSLAPIILDVLAEISTRLOFAEGE
GKFSHFVAITLISLTKAPLVPAQTPVNVNLSQRMLNLEIILACVGLAPENMMILE
YK"

ALIGNMENT Scores:
Pred. No.: 8,67e-184 Length: 1538
Score: 2400.00 Matches: 456
Percent Similarity: 95.69% Conservative: 32
Best Local Similarity: 89.41% Mismatches: 22
Query Match: 91.22% Indels: 0
Db: 8 Gaps: 0
US-10-718-952-12 (1-510) x AF120146 (1-1538)

QY 1 MetPhe11eGluAenPhe1yVal1GluSerProAenVal1ySerThrGluThrGlu1ue 20
 Db 1 ATGTTATGAGAGGCTTCAAGTTGAGAGCCCAAGCTGAATGACACAGGAACGAGATC 60
 QY 21 GlnSerVal1ySerAntyrg1uThrGluLeuVal1H1eGluAenPhe1yVal1yThrTy 40
 Db 61 CATTCGTGTACGATTACGAAACCAAGGAGTGTCTCATAGAAACCGTAAACGCACTTAC 120
 QY 41 GlnTP11eVal1ySerPro1ySerVal1Antyrg1uPhe1yThrAenThi1sValPro 60
 Db 121 CAATGGGTGCTCAAGCCCAAAACCGTCAATACGATTTCAAAACACACACTCCGCTCCC 180
 QY 61 LysLeuG1yVal1MetLeuVal1G1yTPG1yG1yAenAenG1ySerThre1uThrG1yG1y 80
 Db 181 AACCTGGGGCTTATGCTCGTTGATGGAGGAAACATGATCAACCTTAAACCGCTGGA 240
 QY 81 Val11eAlaAenAenG1uAenP11eSerTP11eAlaThre1yAenPhe1yVal1eG1uAen 100
 Db 241 GTGATCGCCCAACAAAGAGGATCTCGTGGCGACTTAAGACAAAGTGAACAGCAAC 300
 QY 101 TyrPheG1ySerLeuThrG1uA1eSerAla1eArgVal1G1ySerPheG1uG1uG1u 120
 Db 301 TACTTGGCTCCCTCACTCAAGCTTCTTCGATCCGGTGGGATCTTACATAGTGTGAAG 360
 QY 121 11eTyAlaProPhe1ySerLeuLeuProMetVal1AenProAenPhe1yVal1PheG1y 140
 Db 361 ATCTACGCTCTTTCAAAGACCTCTCCCATGATGTAACCTGACCAATGTGTGTTGGT 420
 QY 141 G1yTP11eA1eSerAenAenAenA1eAenA1eAenA1eAenA1eAenA1eAenA1eAen 160
 Db 421 GGTGGGACATTAAGCAGATGAATTAAGACAGCCTATGGCTTAAGACAGCAAGTCTTGAAC 480
 QY 161 11eAenPheG1uAenG1uAenAenAenAenAenAenAenAenAenAenAenAenAen 180
 Db 481 ATTGATCTCCAGAAACAGCTCAGCCGCTACATGACACATGCTCCACCTCCGCTGATAC 540
 QY 181 TyrAenProAenPhe11eAlaAenAenG1uAenAenAenAenAenAenAenAenAenAen 200
 Db 541 TACGACCTGACTTCAACGCTGCGAATCAAGGGTCAAGGCTCAATAGCCGATCAAGG 600
 QY 201 Thr1yG1uAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAen 220
 Db 601 ACCAAGAAAGAAACAAATCGACCGATCAATCAAGACATGAAGGAGTTTAAAGAGAAAGAAC 660
 QY 221 LysVal1AenPhe1yVal1Val1e1uTP11eA1eAenThre1yG1ySerAenAenAen 240
 Db 661 AAGGTGATTAAGCTGCTGTTCTGTGAGACGCAACACAGAGGCTTAACAGCATGGTGC 720
 QY 241 Val1G1yLeuAenAenPhe1yMetG1uAenAenAenAenAenAenAenAenAenAenAen 260
 Db 721 GTAGGGCTTAAACGACGATGAGAAATCAATGGCTTCTGTGAAGAAAGAGAGTGTGAG 780
 QY 261 11eSerProSerThre1yThra1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA 280
 Db 781 ACTGCGCTTCAACGCTATATGAGATTGCTGTGCTTGAAGGCACTTCTTCAATTAAT 840
 QY 281 G1ySerProG1uAenThre1yVal1ProG1yLeu11eA1eA1eA1eA1eA1eA1eA1eA 300
 Db 841 GGGAGTCCCGACAGAACGTTTGTCTGAGCATCATGAGCTGGCTATATCAAGAACTGT 900
 QY 301 Leu11eG1yG1yAenAenPhe1ySerG1yG1uThre1yMet1ySerVal1e1uA1eA1e 320
 Db 901 TTGATCGAGTGGGATACATTCAGAGTGTGACATCAAGTGAAGTCTGTTTGTGTTGAT 960
 QY 321 Phe1eVal1G1yA1eA1y11e1yP11e1yP11e1yA1e1e1yA1e1e1yA1e1e1yA1e 340
 Db 961 TTCCTTGTGGAGCTGGATCAAGCTACTGATTTGTGAGCTATATCATTTGGGAAAC 1020
 QY 341 AenAenAenG1yMetAenAenAenA1eProG1uThre1yPhe1ySer1yG1u1eSer1ySer 360
 Db 1021 AATGACGGGATGAACCTCTCAGCGCTCAAGACTTTTATGATCTTAAGAGATCTCAAGAGAC 1080
 QY 361 AenVal1A1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA 380

Db 1081 AATGTTGTGAGACACATGCTGCTACTTAACGATACCTCTTGTAGCTGTGTAACCTCT 1140
 QY 381 AenPhe1Val1Val11e1ySerTyVal1ProTyVal1G1yAenPhe1yVal1AenAenAen 400
 Db 1141 GACCATGTTGTGCTATCAAGATATGTTCTTATGTTGGGATATGTAAGAGAGCCATGAGAC 1200
 QY 401 G1uTyThre1ySerG1u1ePhe1yMetG1yG1yAenThre1yVal1e1uA1eA1eA1eA1e 420
 Db 1201 GAGTACACCTCGAGATATTCATGAGGAGGAAAGAAATACATTTGATATGATACACTTGT 1260
 QY 421 G1uAenPhe1yLeuA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1e 440
 Db 1261 GAGATCTCTTGTGCTGCTCTCAATCACTTGCATCTTGTCTTCTCCCTGAGCTCAAC 1320
 QY 441 ThrArg1eG1uPhe1yVal1A1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1e 460
 Db 1321 ACCAGAAATCAATTCACAACTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
 QY 461 Thr11eAenPhe1yThre1yThre1yThre1yThre1yThre1yThre1yThre1yThre 480
 Db 1381 ACCATACCTACATTAACCTCAAAAGACCTCTTGAACAGCCGAGACACTGTGTGAAC 1440
 QY 481 A1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1e 500
 Db 1441 GCTCTATTAAGCAGAGGCTATGCTGAGAAACATCTCAAGGCTGCTGCTGCTGCTGCTG 1500
 QY 501 ProG1uAenAenA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1eA1e 510
 Db 1501 CCAGAAACACATATGATTTGGAATACAG 1530
 RESULT 12
 AF120147 1538 bp mRNA linear PLN 23-APR-1999
 LOCUS Tr1c1c1u1m a1c1l1v1ar Taber myo-1n1s1t1ol 1-1ph1s1p1ate s1n1th1ase
 DEFINITION (MIPS) mRNA, complete cds.
 ACCESSION AF120147
 VERSION AF120147.1 GI:4589063
 KEYWORDS Tr1c1c1u1m a1c1l1v1ar (bread wheat)
 SOURCE Tr1c1c1u1m a1c1l1v1ar
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Tr1c1c1eae; Tr1c1c1u1m.
 REFERENCE 1 (bases 1 to 1538)
 AUTHORS Huseain,A., Yan,W., Bock,C., Baga,M., Chibbar,R. and Georges,F.
 TITLE cDNA clone for myo-1n1s1t1ol 1-1ph1s1p1ate s1n1th1ase from wheat
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1538)
 AUTHORS Huseain,A., Yan,W., Bock,C., Baga,M., Chibbar,R. and Georges,F.
 TITLE Direct Subm1s1on
 JOURNAL Submitted (14-JAN-1999) National Research Council Canada, Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada
 FEATURES
 source
 1..1538
 /organism="Tr1c1c1u1m a1c1l1v1ar"
 /mol_type="mRNA"
 /cultivar="Taber"
 /db_xref="taxon:4565"
 gene
 1..1538
 /gene="MIPS"
 CDS
 1..1533
 /gene="MIPS"
 /EC_number="5.5.1.4"
 /function="converts glucose 6-phosphate into myo-1n1s1t1ol 1-ph1s1p1ate"
 /note="cytosolic enzyme"
 /codon_start=1
 /product="myo-1n1s1t1ol 1-ph1s1p1ate s1n1th1ase"
 /protein_id="AAD26331.1"
 /db_xref="GI:4589064"
 /translation="MFIESFKVESFNVKTEHHSVDYETTEVVENRNCTQMVV

KKRTVKYDFKTDTRVPLKGLVMLVGWGNNGSITLITAGVIANKEGISMATKXVQOANYF
 GSLTOASSIRYGVSNYGBEIIYAFPSKILPMVNPDDVYFGWMDISDNILADIKOMREFEK
 IDLOKOLRPMYHMYBLPCIYDPDFKILPMVNPDDVYFGWMDISDNILADIKOMREFEK
 KRKVDKIVLMTANTERYSDVYVGNIDMENTLASEVENKESLISPTLITAIACVLEGI
 PRINSRPOUTPYRGLIELAISKNCILGSDFRSGOTRKSUSVVDLVNAGIKRPSIVS
 YNHLNNDNDNLSAPOTERKSISKNVDDVANSNGILFEGEHPDHVVKIYPIYV
 GDSKAMDEYTSIEIFMGKNNTVMANTCDSILAAPIILDIVLLELSTRLOPKAGE
 GKFHSFHPVATILSYLTKAPLVPAGTIPVVALNSKORAMENILRACVGLAPENNMLE
 YK"

ORIGIN

Alignment Scores:
 Pred. No.: 8, 67e-184 Length: 1538
 Score: 2400.00 Matches: 456
 Percent Similarity: 95.69% Conservative: 32
 Best Local Similarity: 89.41% Mismatches: 22
 Query Match: 91.22% Indels: 0
 DB: 8 Gaps: 0

US-10-718-952-12 (1-510) x AF120147 (1-1538)

QY 1 MetPheIIeGluAsnPhelYserValGluSerProAsnValIleTyrThrGluThrGluIle 20
 Db 1 ATGTTTCATGAGAGCTTCAAGAGTGAAGCCCAACGTGAAGTACACAGAGAAACGAGATC 60
 QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValIleGluAsnArgAsnGlyThrTyr 40
 Db 61 CATTCGTCTGATGATTACGAAACACGAGAGTCTGTCATGAGAACCGCTAACCGCACTTAC 120
 QY 41 GlnTrrPLeValIleYserProIleYserValAsnTyrGlnPhelYserThrAsnThrHiValPro 60
 Db 121 CATGAGTGTGTCAGAGCCCAAAACCGTCAATATGATTCATAAAGACAGACGCGTCCCC 180
 QY 61 LysLeuGlyValMetLeuValGlyTrrPGLYGLYAsnAsnGlySerThrLeuThrGlyGly 80
 Db 181 AAGCTAAGGCGTATGCTCGTGGATGGGAGGAAACATGATCAACCTTAAACCGCTGSA 240
 QY 81 ValIleAlaAsnArgGluAsnIleSerTrrPALAThrLysAspLysIleGlnGlnAlaAsn 100
 Db 241 GTGATGCGCAACAAGAGGATCTCTGGCGGACCTAAAGACAAAGCTGCAACAGCAAAAC 300
 QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
 Db 301 TACTTTGGCTCCCTCACTCAAGCTTCTTGATCCGGGTGGATCTTACATATGATGTAAG 360
 QY 121 IleTrrAlaProPheYserLeuLeuPrometValAsnProAspAspIleValPheGly 140
 Db 361 ATCTAGCTCTCTTCAAGAGCTCTCCCATGGTGAACCTGACGATGTGTGTGGT 420
 QY 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 Db 421 GGTGGGACATTAAGCAATGAACCTTACGACAGCGCATGGCTAGACCAAGATCCTTGAC 480
 QY 161 IleAspLeuGlnIleGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
 Db 481 ATTGATCTCCAGAAACAGCTCAGGCGGTACATGACCAATGGTCCACTCCCTGGTATC 540
 QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluIleArgAlaAsnAsnValIleLysGly 200
 Db 541 TACGACCTGACTCTATCCCTGCGAATCAAGGCTCACGTGCTAATAGCTGATCAAAAGGT 600
 QY 201 ThrLysGlnGlnIleGlnIleIleLysAspIleLysAspIleLysAspIleLysAsp 220
 Db 601 ACCAAGAAAGAAACATGACCAACGATCATCAAGATAGAGAGAGTTTAAGAGAAAGAAC 660
 QY 221 LysValAspLysValValIleuTrrPrrAlaAsnThrGluArgTyrSerAsnLeuVal 240
 Db 661 AAGGTGATTAAGCTGTGTCTGTGAGACAGCAACACAGAGGTTACAGCGTGGTGC 720
 QY 241 ValGlyLeuAsnAspThrMetGluuAsnLeuValAlaValAspArgGlnGluValGlu 260
 Db 721 GTAGGGCTTAACGACGATGAGGAATCTATTGCTGTGAGAAAGACGATCTGAG 780

QY 261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
 Db 781 ATCTGCGCTTACACGCTATATGCGATTCCTGTGTCTTGAAGGATCTCTTCACTTAT 840
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 Db 841 GGAATCCGCGAAGACCTTTGTCTTCTGACTCATGAGCTGGCTGATATCGAAGACGTGT 900
 QY 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
 Db 901 TTGATCGGTGGGAGATGACTTCAAGAGTGTGAGACTAATGATGAATGCTGTTTGGTTAT 960
 QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
 Db 961 TTCCTTGTGAGAGCTGGCATCAAGCCCTCTTCAATGTGAGCTAATATTCATTGGGAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerLysProGlnThrPheArgSerLysGluIleSerLysSer 360
 Db 1021 AATGACGCGATGAAACCTTCAGCGCTCAGACTTTAATCTTACGAGATCTTCAAGAGAC 1080
 QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
 Db 1081 AATGTTGTGAGACATGCTGGCTAGTAACGCTATCCTTTGAGCTGGTGAACATCT 1140
 QY 381 AspHisValValIleLysTrrValProTyrValGlyAspSerLysArgAlaMetAsp 400
 Db 1141 GACCATGTTGTGCTCATCAAGATATGCTTATGTTGGGATAGTAAGAGACCATGAGAC 1200
 QY 401 GlyTrrThrSerGluIlePheMetGlyLysAsnTrrIleValLeuHisAsnThrCys 420
 Db 1201 GAGTACACCTCGAATATTCATGAGGAGGAAAGATGATTTATCATTAACACTTGT 1260
 QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
 Db 1261 GAGGATCTCTTTTGGCTGCTCCATCATCTTGATCTTGTCTTCTGCTGAGCTGACG 1320
 QY 441 ThrArgIleGlnPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
 Db 1321 ACCAAGATTCATTCATAAGCTGAAGGAGGGAAGTTTCATTTCTTCCACCACATACCA 1380
 QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
 Db 1381 ACCATACCTCAGTTACTCAAAAGGACCTCTTGTACCAAGCCGGAGACCTGTGGTGAAC 1440
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluLeuAsnIleMetArgAlaCysValGlyLeuAla 500
 Db 1441 GCTTATCTTAACAGAGGCGTATGCTGAGAAACATCTCAGGCGCTCGCTTGGCTTGG 1500
 QY 501 ProGluAsnAsnMetIleLeuGluTrrLys 510
 Db 1501 CCAGAGAACCAATGATCTTGAATATCAAG 1530

RESULT 13
 AF120148 1538 bp mRNA linear PLN 23-APR-1999
 LOCUS Triticum aestivum cultivar Biggar myo-inositol 1-phosphate synthase
 DEFINITION (MPS) mRNA, complete cds.
 ACCESSION AF120148
 VERSION AF120148.1 GI:4589065
 KEYWORDS
 SOURCE
 ORGANISM
 Triticum aestivum (bread wheat)
 Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticaceae; Triticum.
 REFERENCE
 1 (bases 1 to 1538)
 Huseain, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
 cDNA clone for myo-inositol 1-phosphate synthase from wheat
 JOURNAL
 Unpublished
 2 (bases 1 to 1538)
 Huseain, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
 Direct Submission
 JOURNAL
 Submitted (14-JAN-1999) National Research Council Canada, Plant

Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada

FEATURES

source

Location/Qualifiers

1..1538

/organism="Triticum aestivum"

/mol_type="mrna"

/cultivar="Biggar"

/db_xref="taxon:4565"

1..1538

/gene="MIPS"

1..1533

/gene="MIPS"

/EC_number="5.5.1.4"

/function="converts glucose 6-phosphate into myo-inositol 1-phosphate"

/note="cytosolic enzyme"

/codon_start=1

/product="myo-inositol 1-phosphate synthase"

/protein_id="AAD26332.1"

/db_xref="GI:4589066"

/translation="MPTESPKVSPNVKKTENEHSVDYETTEVVENKNGTQKVV

KRTVYKDKTDTPVKLGMVWGNGNSTLTAGVANKGISWATKXQOANVF

GSLTQASSIRVGSYNGEIEIYAPFKSLPMVNPDPVFGGDIIDMNLADAMARVLD

IDLOKLRMEHMYPLPGIYDPDPIAAGSRANSVIGTKKEQVDOIIKMRERKE

KKVDKLVLTANTERYSDDVVGLNDTMENTLASEKEKESISPTLYAIACVIGI

PRINSRQPTFVPGLELAISKCLIGDDPKSGOTKMSVVDPLVGAGIPTSIVS

YNHLGNDKMLSAIPOTFRSKISKNVDDVAVNGILFEGSEPHDVIVKYPVY

GSKRAMDYTSIEIFMGGNITVMNHTCDLAAPIIDLVLAELSRVIVKPKSGE

GKFSHPHVAITILSYLTAKAPLVPAQTPVNVNLSKORAMENILRAVCGLVAPNNMILE

YK"

ORIGIN

Alignment Scores:

Pred. No. : 8.67e-184

Score: 2400.00

Percent Similarity: 95.69%

Best Local Similarity: 89.41%

Query Match: 91.22%

DB: 8

US-10-718-952-12 (1-510) x AP120148 (1-1538)

QY 1 MetPheilegluaasPhelysValgiuserProaenVallyeThrgiunThrgiunle 20

DB 1 ATGTCATCGAAGGCTCAAGGTGAGACCCCAAGAGTAAAGTACCAAGACGAGATC 60

QY 21 GlnserVallyeAsnTyrgiunThrgiunleValhieguaasnglyThryr 40

DB 61 CATTCGTGTACGATTACGAACACGAGGCTGCTCAAGAACCGCTAACCGCACTTAC 120

QY 41 GlnTPRileVallyeProlyserValaenTyrginPhelyeThraanthriValPro 60

DB 121 CAATGGGCTGTCAGACCCCAACCGTCAATATGATTCAAAACAGACTGCGGCCCC 180

QY 61 LysleuValMetleuValgiyTPRgiyuaasnglyserThreuthrgiunly 80

DB 181 AAGCTAGGGGTATGCTGTTGATGAGGAGAAACAATGATCAACCTTAAACCGCTGGA 240

QY 81 ValilealAasnrgluaspiileserTPRAThrlyAspLyellegingluAaen 100

DB 241 GGTATCGCCAAAGAAAGGATCTCTGGGCACTAAAGACAAAGTGCACACGAAAC 300

QY 101 TyrPhegiyserleuthrgiunAaseralailaerqValgiyserPhegiinglyugiu 120

DB 301 TACTTGGGCTCCCTCACTCAAGCTTCTTGATCCGGGTGGATCTTAAATGATGTAAG 360

QY 121 IleTyralaProPheylserleuThreuthrgiunleValaenProaaspriyleValPhegiy 140

DB 361 ATCTAGGCTCTTCAAGGCTCTCTCCCAAGTGAACCCCTGACATGATGTGTGGT 420

QY 141 GlyTPRaspriileserAasnMetAasnleualAaspaalAmetAlaargAlaValPheap 160

DB 421 GGTGGGACATAGGACATGAACCTTACGAGACGATGCTAGAGCAAGATCTCTTGAC 480

QY 161 IleaspleuGlnlyeGlnleuargProTyrmecGlyuserMetValProleuProgiyle 180

DB 481 ATTGATCTTCAAAAACAGCTCAGCGCTACATGAGCACATGCTCCACTCTCCGTATAC 540

QY 181 TyrAaspProaaspPheilealAaenngingluAargAlAasnValIlelysgly 200

DB 541 TACGACCTGACTTATATGCTGCAATCAAGGTCACGTGCTTAATAGCTGATCAAAAGT 600

QY 201 ThrlysgingluGlnValglnGlnIleIlelysaaspiilelysaalAphelysgluAthr 220

DB 601 ACCAAGAAAGAACAGATGACACAGATCATCAAAAGCATGAGGAGATTAAAGAAAGAAC 660

QY 221 LysValaaspleyValValleuthrPThrAlaanthrgiunAargTyserAasnleuVal 240

DB 661 AAGGGATACCTGATGCTGTTGTCGACAGCCACACAGACGTTACGACGATGTGGTC 720

QY 241 ValGlyLeuAasnAspThrmecGluAasnleuAlaAalAaAspAargngluAgiu 260

DB 721 GTAGGGCTTAACGACACATGAGAGANTCTATTGGCTTCTGTGAGAGAACAGTCTGAG 780

QY 261 IleserProserThreuthrgiunAalAcyValMetGluAasnValProPheileasn 280

DB 781 ATCTGCGCTTACGCTATATGCAATGCTGTTGTTGTTGAAGGATTCCTTCAATTAAT 840

QY 281 GlyserProGlnAasnThPheValProgiyleuIleaspleuAlaIleAlaAarganthr 300

DB 841 GGGATCTCCGACAAACAGTGTGCTTCTGACATCAGCTGCTGCTATATCGAAGAACGT 900

QY 301 LeuilegLyAspAaspPheylserGlyGlnThryrlyMetlyserValleuValaap 320

DB 901 TTGATCGGTGGGAGATGACTTCAAGAGTGTCTCAGACTTAAGATAGTCTGTGGTGTAT 960

QY 321 PheleuValgiyAlaGlyIlelyserProThrseryleValserTyraasniIseuGlyAaen 340

DB 961 TTCCTTGTGCGAGCTGCGGATCAAGCTTCTGATGTGAGCTTAATCATTTGGGAAC 1020

QY 341 AasnAspGlyMetAasnleuSerAlaProGlnThrPheAargserlyGluIlelyser 360

DB 1021 AATGACGCGATGAACTCTCAGCGCTCAGACTTTTAAATTCAGAGATTCCTCAAGAGAC 1080

QY 361 AasnValAaaspAaspMetValAasnSerAaAlaIleuTyrgluProgiyGluIleuPro 380

DB 1081 AATGTTGTGACAGATGAGTGTGCTAGTAACGTTATCTCTTGACCTGTGTAACATCT 1140

QY 381 AspiIseValValIlelyserTyraProTyraValgiyAaspserlyeAargAlaMetAap 400

DB 1141 GACCATGTTGCTGATCATGATATGTTCTTATGTTGGGATAGTAAGAGACCATGAGAC 1200

QY 401 GlnTyThrseryGluIlePheMetGlyGlyAasnThriIleValleuHiAasnThry 420

DB 1201 GAGTACACCTCGAAGATATTCATGAGGAGAAAGAAATAGCATTTGATGATCACTTGT 1260

QY 421 GluaaspleuLeuAlaProIleIleuAaspleuValleuAaIleuAaIleuAaIleuAa 440

DB 1261 GAGGATCTCTTGTGCTGCTCAATCAATCTTGATCTTGTCTTGTGCTGCTGACGCT 1320

QY 441 ThraArgileGlnPheylAaGluAasnGlylyPheIseSerPhehiAproValAla 460

DB 1321 ACCAAGATTCATTAAGGCTGAAGGAGGAGGAGGAGTTCTTCCACACCAATAGGA 1380

QY 461 ThrIleleuserTyreuthrgiunAaProleuValProProgiyThProValValaen 480

DB 1381 ACCATATCTCAGTTATCTCAAAAGGACATCTTGTATCCAGCGGAGACCTGTGTGTAAC 1440

QY 481 AlaIseSerlyGlnAargAlaMetleuGluAasnIleMetAargAlaCysValGlyleuAla 500

DB 1441 GCTTATCTTAAGCAAGGCTATGCTGAGAAACATCTCAAGGCTGCTGTGGGCTGCTG 1500

QY 501 ProGluAasnMetIleleuGluTyryls 510

DB 1501 CCAGAGAACACATGATCTTGGAAATACAG 1530

RESULT 14
 AF433879 1986 bp mRNA linear PLN 07-NOV-2001
 LOCUS Sueda maritima subsp. salsa myo-inositol-1-phosphate synthase
 DEFINITION (INPS) mRNA, complete cds.
 ACCESSION AF433879
 VERSION AF433879.1 GI:16755888
 KEYWORDS
 SOURCE
 ORGANISM
 Sueda salsa
 Sueda salsa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Amaranthaceae; Sueda.
 1 (bases 1 to 1986)
 Wang, J.P., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.
 Direct Submission
 Submitted (15-OCT-2001) The Biology Department of Shandong Normal
 University, Key Laboratory of Plant Stress Research, No.88, Wenhua
 East Road, Jinan, Shandong Province 250014, China
 FEATURES
 source
 1..1986
 /organism="Sueda salsa"
 /mol_type="mRNA"
 /sub_species="salsa"
 /db_xref="taxon:126914"
 1..1986
 /gene="INPS"
 /gene="INPS"
 140..1672
 /gene="INPS"
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase"
 /protein_id="AAU28131.1"
 /translation="MFIEENFESPNVNTENEIHSVNYETTELVHENRKGSGYOMI
 VKPPTVKYQFQTHVPGLGVMLVWGSGNNECTLTGVIANREGISWATKDIQANVY
 FGSITQASIRVGSFNGBEIYAPFSILPVPNPDIYEGWDIDLMADMTAKVTLV
 DIDLOKORPYMEHWPRLPGYVDDPFLAANDSSANNITKGTKEQYEOVYKIDREK
 EKKAKDKVLTMTANTERYISVVVGLDNTENLMAAYRNESEISPSVIAIAGIED
 VPFINGSPQNTVFPGLLEIAIKKNCILGGDFFKSGQTKMSLVDPFLVGAKIKPTSLV
 SYNHNQNDGNMLNAPQTFERSKEISKSVVDVAASNGIILPEBHPDHVAVIKYVY
 VGSKRAMDVEYTSIEFMGKNTIYLHNTCEBSLLAAPILDLVLAELSTIEFKAPT
 EDKHTTFRPVATILISYLKAPLVPGRFVVNALSKQPMLEKMACVGLAPENNMLE
 YK"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,436-183 Length: 1986
 Score: 2399.00 Matches: 462
 Percent Similarity: 94.72% Conservative: 22
 Best Local Similarity: 90.41% Mismatches: 25
 Query Match: 91.18% Indels: 2
 Gaps: 2
 US-10-718-952-12 (1-510) x AF433879 (1-1986)
 QY 1 MetPheIleGluAenPheIysValGluSerProAenValIySYrThhGluThhGluIle 20
 Db 140 ATGTCATCGAAATTTCAAGGTAAAGACCCCAATGTTAACTACACGAAATGCAAAATT 199
 QY 21 GlnSerValIYrAsnTYrGluThhThrChhGluLeuValHISGluAenArgAenGlyYThr--- 39
 Db 200 CATTAGTTTCAATTAATGAAACCACTGAATTAAGTTTCACGGAATCGCAAAAGGCTTGCT 259
 QY 40 TYrGlnTPILeValIysProIysSerValAenTYrGlnPheIySThAsnThhThIsvaI 59
 Db 260 TATCATGATGATCGTCAAACTTAAACCGTCAAAATTCACCAATTCGCAAACTCATGAT 319
 QY 60 ProIysLeuGlyValMetLeuValGlyTYrPGLyGlyAenAenGlySeThrLeuThhGly 79
 Db 320 CCAAACTTGAGGATATGTTGGTGGGAATGGGGTGGCAACACGATTGACACACTACCGGT 379
 QY 80 GlyValIleAlaAenArgGluAspTLeSerTTPaATThrIyAspIyIleGlnGlnAla 99

Db	380	GGTGTCAATCGCAATCGCAGGGATATATCATGGGCAACAAAGACAATTAATCAACAGCT	439
OY	100	AsnTyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu	119
Db	440	AACATATTTGGATCCCTCACCTCAGCGCTTCCTTATGTGTGGTCTTTTAAATGAGAG	499
OY	120	GluIleTyrAlaProPheGlySerLeuLeuProMetValAsnProAspAlaValPhe	139
Db	500	GAGATTATGCTCTCTTTAAGACCTTCTTCTTATGGTGAACCAAGATGATATGCTTT	559
OY	140	GlyGlyTyrAspAlaSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaValPhe	159
Db	560	GGGGGCTGGGACATTAAGTGAAGCTTGAACCTAGCAGACGCAAGCAAGAGGCAAGGTGTG	619
OY	160	AspIleAsnLeuGlnGlyGlnLeuArgProTyrMetGluSerMetValProLeuProGly	179
Db	620	GACATTGATTTGGCAAAAGCAGTTGAGACCTTACATGAGCACAATGTTCCCTTCGTGT	679
OY	180	IleTyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLeu	199
Db	680	ATCTATGACCTGATTTTCATTTGCTGCTTACCAACATTTCTGTCGAAATACATTTAAG	739
OY	200	GlyThrArgGlnGlnGlnValGlnGlnIleIleLeuAspIleLeuValaPheGlyAla	219
Db	740	GGCACTAAGAAAGAGCAAGTGAACAAAGTCAATTAAGACATCAGGAATTTAAGAGAG	799
OY	220	ThrIysValAspIysValValIleuITrThrAlaAsnThrGluArgYrSerAsnLeu	239
Db	800	AACAAAGTGGCAAGGTGTGTCTCTGTGCAAGCAAAACAAGAGAGTACAGTATGTG	859
OY	240	ValValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValaAspArgAsnGluAla	259
Db	860	GTTGTGGACTCAATGACACCAACAGACACTTATGGCAGCTGTGGAAGGAATGAGCT	919
OY	260	GluIleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIle	279
Db	920	GAGATTCTCCATGACTTGTGTATGTATGCTATGCTTGATCCCAAGAGATGTCCCTTCATC	979
OY	280	AsnGlySerProGlnAsnThrPheValProGlyLeuIleAsnLeuAlaIleAlaArgAsn	299
Db	980	AACGTATGTCCTCGAACAACCTTGTCTCTGTGTCAATGGAACGGCATTAAGAAAGAT	1039
OY	300	ThrLeuIleGlyGlyAspAspPheLeuSerGlyGlnThrIysMetClySerValLeuVal	319
Db	1040	TGTTTATGAGGAGATGACTTCAAGATGTGTCAACTTAAGATGAATATCAGTCTCTT	1099
OY	320	AspPheLeuValGlyAlaGlyIleLeuProThrSerIleValSerTyrAsnIleLeuGly	339
Db	1100	GACTTCTCTGTTGGGGCTGTATTAAGCAACATCAATGTGTGAGCTCAACACACTTGGGA	1159
OY	340	AsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerGlyGluIleSerIys	359
Db	1160	AACATATGATGGAGTGAACCTTTCAGCACTTCGACCTTCAGTCTTAAGAAATTTCAAG	1219
OY	360	SerLeuValValaAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHis	379
Db	1220	AGTATATGTGTGATGACATGTGTCTAGCAACAGGTATCTTGATAGACCTGTGTGAACAC	1279
OY	380	ProAspHisValValIleLeuTyrValProTyrValGlyAspSerIysArgAlaMet	399
Db	1280	CTGTATCATGTGTGTTCATCAATATGTTCATATGTCCGAGACAGGAAGAGCTATG	1339
OY	400	AspGlyTyrThrSerGlyIlePheMetGlyGlyIysAsnThrIleValLeuHisAsnThr	419
Db	1340	GATGAGTACATCTGAGATATATCATGGTGTGTGAAGAACACATTCGTTTGTGCAACAACC	1399
OY	420	CysGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleLeuAlaGluLeu	439
Db	1400	TGTAGAGACTCTCTCTGGCGCTCCATCATCTTAAGCTTGCTTACCTTGGCCGAGCTT	1459
OY	440	SerThrArgIleGlnPheIysAlaGluAsnGlnIlyIysPheHisSerPheHisProVal	459
Db	1460	AGCAACCGCATTTGAATTCAAAGCCGGAACCTGAGACAAAGTTCCACACCTTCATCTCAGTG	1519

QY	460	1AATHTTILLeuSeTtYleuThrYvalaPProdeuValProProGlyTThProvalVal	479		
Db	1520	GCTACCAACTCAAGTTATCTAATCAAAACACCTCTGGTTCCACCAAGTACCTCCGCTGCTG	1579		
QY	480	AenAlaLeuSeTtYGLnArgAlaMetLeuGluAenIleMetArgAlaCysValGlyLeu	499		
Db	1580	AATGGCTCTCAAAACACAGGCGCATCTTAGAGAAA---ATGAGGCGATGTGTAGTTTA	1636		
QY	500	AlAProGluAenAMetIleLeuGlyTtYrIys	510		
Db	1637	GCTCAGAGAACAAACATGATCCTTAGTACAAG	1669		
RESULT 15.					
MCU32511		2053 bp	mRNA	linear	PLN 11-JUN-1996
LOCUS					
DEFINITION		Meesembryanthemum crystallinum myo-inositol-1-phosphate synthase			
ACCESSION		U32511			
VERSION		U32511.1			
KEYWORDS		GI:975887			
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCES					
AUTHORS					
TITLE					

[illegible]

Db 1262 AAAGCAACGTGTCATGACATGGTTGCCAGCAACGGCATCCTCTATGAGCCGGGCGAA 1321
Qy 379 HisProAspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAla 398
Db 1322 CACCCCTGATCATGTGTGTGTCATCAAGTATGTGCCATATGTGGAGACAGCAAGAGAGCC 1381
Qy 399 MetLaspGluTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuHisAsn 418
Db 1382 ATGGATGAGTACACTTCGGAATAATTATGGGGGGGAAACGACACTATGTCTCATGCACAAC 1441
Qy 419 ThrCysGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuLeuAlaGlu 438
Db 1442 ACCGTGAGGACTCTCTCTTGGCTGCCCCCATCATCTGACTTGGTCTCTCTGCTGAA 1501
Qy 439 LeuSerThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisPro 458
Db 1502 CTAAAGCAACCCGATTCAGCTGAAGGCTGAAGGAGAGGACAACTTCCACTCTCCATCCG 1561
Qy 459 ValAlaThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProVal 478
Db 1562 GTGGCTACCATCTGTGTTACCTCAACAGGCCCTCTTGTTCACCAAGGCACCTCCAGTG 1621
Qy 479 ValAsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGly 498
Db 1622 GTAATTCACCTTCAGAGCAGAGGCTATGCTGAGAACATACATGAGGGCTTGTGTGGT 1681
Qy 499 LeuAlaProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1682 TTGGCTCCTGAGAACACATGATCCTCGAATACAAAG 1717

Search completed: June 7, 2005, 20:45:40
Job time : 5355 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:10:41 ; Search time 603.5 Seconds

(without alignments)
4765.695 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631
Sequence: 1 MFIEFKYSPVVKXTETBI.....NIMRACVGLAPENNMLEYK 510

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame.p2n.model -DBV=xlp
-O=/cgn2.1/USPTO.spool.p/US10718952/runcat.06062005.173400.12840/app.query.fasta.1.1294
-DB=N.GeneSeq.16Dec04 -QFMT=fasta -SUFFIX=ep2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=humand40.cdi
-LIST=45 -DOCAIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952@cgn.1.1.886@runat.06062005.173400.12840 -NCPU=6 -ICPU=3
-NO.MAP -LARGEOQUERY -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.GeneSeq.16Dec04:*
1: geneeqn1980s:*
2: geneeqn1990s:*
3: geneeqn2000s:*
4: geneeqn2001as:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004as:*
13: geneeqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2631	100.0	1533	12	ADQ14500 Mutant so
2	2631	100.0	1533	13	AD882001 Soybean m
3	2624	99.7	1533	12	ADQ14504 Wild type
4	2624	99.7	1533	12	ADQ14502 Mutant so
5	2624	99.7	1533	13	AD882005 Soybean m

6	2624	99.7	1533	13	AD882003	AD882003 Soybean m
7	2593	98.6	1533	12	ADQ14498	Adq14498 Mutant so
8	2593	98.6	1533	12	AD881999	Ad881999 Soybean m
9	2593	98.6	1533	13	AD881993	Ad881993 Soybean m
10	2593	98.6	1760	12	ADQ14490	Adq14490 Wild type
11	2593	98.6	1782	2	AAV62440	AAV62440 Soybean w
12	2588	98.4	1533	12	AAV62443	AAV62443 Soybean m
13	2588	98.4	1533	12	ADQ14494	Adq14494 Mutant so
14	2588	98.4	1533	13	AD881997	Ad881997 Soybean m
15	2470	93.9	1950	2	AAx90402	AAx90402 Nicotiana
16	2384	90.6	1533	6	AB213633	Ab213633 Arabidops
17	2384	90.6	1533	12	ADN73524	Adn73524 Thale cre
18	2384	90.6	1938	3	AAc50242	AAc50242 Arabidops
19	2376	90.3	1781	4	AAc87643	AAc87643 Brassica
20	2354	89.5	1665	2	AAx09006	AAx09006 Inducible
21	2349.5	89.3	1536	3	AAc43428	AAc43428 Arabidops
22	2349.5	89.3	1837	3	AAc34806	AAc34806 Arabidops
23	2346	89.2	1931	2	AAx24407	AAx24407 Maize myo
24	2337	88.8	1959	4	AAc85922	AAc85922 MIP synth
25	2285	86.8	1759	6	ABQ72653	Abq72653 Human MHD
26	2126.5	80.8	1719	3	AAc51095	AAc51095 Arabidops
27	2040	77.5	1536	12	ADP43918	Adp43918 P. coarct
28	1823	69.3	3546	2	AAx24411	AAx24411 Maize myo
29	1822	69.3	3546	2	AAx24410	AAx24410 Maize myo
30	1616	61.4	1772	4	AAH14171	AAH14171 Human cDN
31	1616	61.4	1818	4	AAH15251	AAH15251 Human cDN
32	1616	61.4	1825	6	ABT07176	ABT07176 Human ova
33	1616	61.4	1852	6	AAH47741	AAH47741 Human hM1
34	1616	61.4	1862	4	AAI59232	AAI59232 Human pol
35	1616	61.4	2380	13	ADQ86789	Adq86789 Human tum
36	1616	61.4	2380	13	ADQ85720	Adq85720 Human tum
37	1616	61.4	2380	13	ACN40520	ACN40520 Tumour-as
38	1612.5	61.3	1605	8	ABT19345	ABT19345 Aspergill
39	1608.5	61.1	1605	8	ABT21165	ABT21165 Aspergill
40	1602	60.9	1833	4	AAH47740	AAH47740 Murine m
41	1602	60.9	2280	4	ABU17361	ABU17361 Drosophi
42	1565.5	59.5	1980	12	ADQ00499	Adq00499 Novel hum
43	1565.5	59.5	1980	12	ADN98930	Adn98930 Novel hum
44	1553	59.0	1991	12	ADG73735	Adg73735 Aspergill
45	1546	58.8	1704	8	ABT20567	ABT20567 Aspergill

ALIGNMENTS

RESULT 1
ID ADQ14500 standard; cDNA; 1533 BP.
XX
AC ADQ14500;
XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.
XX
KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
OS Glycine max.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #3"
XX
FN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX

PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00293315.
PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI; 2004-533135/51.
DR P-PSDB; ADQ14501.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX Claim 10; SEQ ID NO 11; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 3.9e-255 Length: 1533
Score: 2631.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-718-952-12 (1-510) x ADQ14500 (1-1533)

QY 1 MetPheIIeGluAsnPhelysValGluSerProAsnValIyTyrThGluThrGluIle 20
DB 1 AAGTTCATCGAATAATTTTAAAGTAGAGTCTCTAATGTGAATACACCGAGCTAGATT 60
QY 21 GlnSerValIyTyrAsnTyrGluThrThrGluLeuValIHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCTGTACAACTACGAAACCAACGAACTGTTCACGAGAACGGAATGGACCTAT 120
QY 41 GlnTPrIleValIyProIySerValAsnTyrGlnIlnPhelysThrAsnThrHisValPro 60
DB 121 CAGTGGATTGTCAAAACCCAAATCCGTCAACTACCAATTTAAACCAACCCATGTTCCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyIAsnAsnGlySerThrIleuThrGlyGly 80
DB 181 AAATTGGGGGATGATCTTGTGGGTTGGGGTGGAAACAACGCGCTCAACCTCCCGTGGT 240
QY 81 ValIleAlaAsnArgGluAsnProIleSerTPrAlaThrIlyAsnIlyIleGlnIlnAlaAsn 100
DB 241 GTTATTGTCAACAGAGGACATTTCAATGGGCTCAACAAAGACAAATTCACCAACCAAT 300

QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTGGCTCCCTCACCAAGCCCTCAGCTATTGAGATTGGATCTTCCAGAGGAGAGAA 360
QY 121 IleTyrAlaProPheIySerLeuLeuProMetValAsnProAsnProIleValPheGly 140
DB 361 ATCTATGCCCATTCAGAGATCTCTCCAAATGGTAAATCTTCAGACATTTGTGGG 420
QY 141 GlyTPrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIyValPheAsp 160
DB 421 GGATGGCATTCACCAACATGAACTGGCTGATGCCATGGCCAGGCAAGATGTTTGAC 480
QY 161 IleAspLeuGlnIyGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTGGCAGAGACAGTTGAGGCTTACATGAAATCCATGGTTCCACTCCCGAATC 540
QY 181 TyrAspProAsnProIleAlaAlaAsnGlnGluIlnArgAlaAsnAsnValIleIyGly 200
DB 541 TACGACCCGAAATTCATTTGCTGCCAACCAAGAGAGCGTGCACACATGATTAAGGAC 600
QY 201 ThrIySGlnGlnIlnValGlnIlnIleIyAspIleIyValIlePheIyGluAlaThr 220
DB 601 ACMAAGCAAGCAAGATTCCAGCAATCATCAAGACATCAAGGGGTTTAAGAAAGCCACC 660
QY 221 LysValAspIyValIleValIleuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGCACAAAGTGGTGTCTGTGACCTGCCAACAGAGAGTATAGCAATTTGGTT 720
QY 241 ValGlyLeuAsnAsnProThrMetGluAsnLeuLeuAlaValIleAspArgAsnGlu 260
DB 721 GTAGGCTTAAATGACACCATGAGAAATCTCTTGCTGTGTGGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCCTCTTCCACTTGTATGCCATTCCTGTGTGATGGAATGTTCTTCATTATAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaIyAsnThr 300
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGGGCTGATGTGCTTGGCCATCGCAGAGAACT 900
QY 301 LeuIleGlyIyAsnAsnProPheIySerGlyIlnThrIySerMetIySerValLeuValAsp 320
DB 901 TTGATGTGTGAAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGGTAT 960
QY 321 PheLeuValGlyAlaGlyIleIyProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTTCTTGGGGGGCTGGATACCAACCAACATCTTATAGTACTTACCAACCACTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIyGlyIleSerIySer 360
DB 1021 AATGATGGTATGAATCTCTCGGCTCCACAAACCTTCGCTCCAGGAATCTCCAAAGAC 1080
QY 361 AsnValValAspAsnProMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisAsp 380
DB 1081 AACGTTGTGACGATAGTGAACAGCAATGCCATCTCTATGAGCCCTGGAGCAATCCC 1140
QY 381 AspHisValValIleIyTyrValIleProTyrValGlyAspSerIyArgAlaMetAsp 400
DB 1141 GACCATGTGTGTATTATTAAGTATGTGCTTACGTAAGGCAATACCAAGGCCATGAT 1200
QY 401 GluTyrThrSerGluIlePheMetGlyIyAsnThrIleValIleuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATTCATGAGGTGGAAGAAACCAATGTTTGGACAAACACATGT 1260
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTAACTGCTCCCTATATCTTGAATCTTCTTCTTGGCTGACCTGAGC 1320
QY 441 ThrArgIleGlnPheIyValAlaGlnGlnGlyIyAspPheHisSerPheHisProValAla 460
DB 1321 ACTAGATCCAGTTTAAAGCTGAAGAAATGAGGAAATTCCTCACTATTCACCCAGTTGCT 1380
QY 461 ThrIleuSerTyrLeuThrIyAlaProIleuValProProGlyThrProValValAsn 480

|||||
DB ACCATTCTAGCTACTGACCAAGGCTCTCTGTTCCACCGGTACACCACTGGTGAT 1440
QY 461 AAlaSerLyGlnArgAlaMetLeuGluAniIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTGTCAAAACAGCGTGCATGCTGGAAACATATAGAGGGCTTGTGTGATTGGCC 1500
QY 501 ProGluAenAenMetIleLeuGlnTyrLys 510
DB 1501 CCAGAGAAATACATGATTTCTCAGTACAAAG 1530
RESULT 2
ADS82001 ID ADS82001 standard; cDNA; 1533 BP.
AC ADS82001;
XX 18-NOV-2004 (first entry)
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
DE
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flautence; mutant.
XX
XX Glycine max; line 29010CP01.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 1. 1533
FT CDS /*tag= a
FT /*product= "myo-inositol 1-phosphate synthase"
FT mutation /*tag= b
FT
FT
PN US2003074685-A1.
XX
XX 17-APR-2003.
PD
XX 11-MAR-2002; 2002US-00025003.
PF
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBA/STAN S A.
PI
PI Hitz WD, Sebastian SA;
XX
XX WPI: 2004-639957/62.
DR P-PSDB; ADS82002.
DR
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phylic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 11; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phylic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of

CC crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phylic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flautence when consumed by humans. The present sequence encodes a mutant myo-inositol 1-phosphate synthase.
XX
XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
Pred. No.: 3.9e-255 Length: 1533
Score: 2631.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-718-952-12 (1-510) x ADS82001 (1-1533)
QY 1 MetPheIleGluAenPheLyValGluSerProAenValLyTyThrGluThrGluIle 20
DB 1 ATGTTTCATCGAAGATTTTAAGGTAGAGAGTCTTAATGTAAGTACACCGAGCTAGATT 60
QY 21 GlnSerValIlyrAenTyrGluThrThrGluLeuValHisGluAenArgAenGlyThrTyr 40
DB 61 CAGTCGGGTCACTACATGAAACCAACCACTGTGTCAAGAAACGAAATGGCACCTAT 120
QY 41 GlnTrpIleValLyPProLySerValAenTyrGlnPheLyThrAenThrHisValPro 60
DB 121 CAGTGGATTGTCACCAACCAACCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAenAenGlySerThrLeuThrGlyGly 80
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGTGGGTGAAACAACGGCTCACTCACTCACTGGT 240
QY 81 ValIleAlaAenArgGluAspIleSerTrpAlaThrLyAspLyIleGlnIleAlaAen 100
DB 241 GTTATTGCTTACAGAGAGACATTTTCATGAGGCTTCAAGACATTCACACACACAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCTCCCTCACCAAGCTCAGCTATTCAGTTGAGTTCCTTCCAGGAGAGAA 360
QY 121 IleTyrAlaProPheLySerLeuLeuProMetValAenProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATGTTATTCGACGACATTTGTTGGG 420
QY 141 GlyTrpAspIleSerAenMetAenLeuAlaAspAlaMetAlaArgAlaLyValPheAsp 160
DB 421 GGATGGATATCAGCAACATGAACTGCTGATGTCATGCGCAGGCGCAAGGCTGTTTAC 480
QY 161 IleAspLeuGlnLyGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTTGCGAAGCACTTGAAGGCTTACATGATGATCATGTTCCCTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAenGlnGluArgAlaAenAenValIleLyGly 200
DB 541 TACGACCGGATTTTCATTTGCTGCGCAACCAAGGAGCGGTGCAACAGTATTAAAGGC 600
QY 201 ThrLyGlnGluGlnValGlnGlnIleIleLyAspIleLyValAenLyGlnAlaThr 220
DB 601 ACAAGCAAGGAGCAAGTTCACCAATCTCAAGACATCAAGCAAGCGGCTTTAAAGACCC 660
QY 221 LysValAlaPheLyValIleValIleLeuThrAlaAenThrGlyArgTyrSerAenLeuVal 240

```

Db 661 AAAGTGA CAAAGTGTGTCTGTGACTGCCA CACAGAGGATATGCAATTTGCTT 720
Qy 241 ValGlyLeuAsnAspThr-MecGluAsnLeuLeuAlaValAspArgAsnGlu 260
Db 721 GTAGCCCTTAATGACACCATGAGAAATCTCTGGCTGCTGTGACAGAAATGAGCTGAG 780
Qy 261 ILeSerProSerThrLeuTyraIleAlaCyValMetGluAsnValProPheIleAsn 280
Db 781 ATTCTCTCCCTTCCACCTTGATGCCATTTGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
Qy 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAACCCCTCAGAACACTTTGTACAGGCTGATTTGATCTTGCCATGCGCGAGAACACT 900
Qy 301 LeuIleGlyValAspAspPheLeuSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTGGGAGAGACTTCACAGAGTGCTCAGACCAGAAATGAAATCTGTGTGTTGAT 960
Qy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyraAsnHisLeuGlyAsn 340
Db 961 TTTCTTGAGGGGCGTGTATCAAGCCAACTATATAGTTACATCAACCATCTGGGAAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db 1021 AATGATGGTATGAATCTCTCGGCTCCACAAACCTTCGCTCCAGAGAAATCTCCAAAGAC 1080
Qy 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyraGluProGlyValHisPro 380
Db 1081 AACGTTGTTGACGATATGCTCAACAGCANTGCCATCTCTATAGACCTGTGTGAAACATCCC 1140
Qy 381 AspHisValValIleLysTyraIleProTyraValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCATGTTGTTGTTATTAATGATATGCTTACGTAGGGGATAGCAAGAGACCATGAT 1200
Qy 401 GluTyThrSerGluIlePheMetGlyLysAsnThrIleValLeuHisAsnThrCys 420
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAGAAACACCATGTTTGCACAAACATGTT 1260
Qy 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGGATTCCTTTTACGCTCTCTTATATCTTGACTTGCTGCTCTTCTGTGAGCTGAGC 1320
Qy 441 ThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAATATTCAGTTTAAAGCTGAAATAGGAGAAATTCATCTCATTCACCCAGTTGCT 1380
Qy 461 ThrIleLeuSerTyThrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGTTCCACCGGATACACCACTGTGAT 1440
Qy 481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCyValGlyLeuAla 500
Db 1441 GCATTGTCAACAGAGGTGCAATGCTGGAAACATTAAGGGCTTGTGTGATTGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuGluTyraLys 510
Db 1501 CCAGAGAAATTAACATGATTTCCAGATACAAAG 1530

```

```

RESULT 3
ADQ14504
ID ADQ14504 standard; cDNA; 1533 BP.
XX AC
XX ADQ14504;
XX
DT 23-SEP-2004 (first entry)
XX
XX Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.
XX
XX Soybean: myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KM inorganic phosphate.
XX
XX Glycine max.
OS

```

```

XX Key Location/Qualifiers
FH CDS 1..1533
FT /*tag= a
FT /product= "Wild type soybean myo-inositol 1-phosphate
FT synthase #2"
PN
XX
XX US2004128713-A1.
PD 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98MO-US006822.
XX 26-APR-1999; 99US-00299315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI; 2004-533135/51.
XX P-PSDB; ADQ14505.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Claim 4; SEQ ID NO 15; 48pp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents cDNA encoding a wild type
XX soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,98e-254 Length: 1533
XX Score: 2624.00 Matches: 509
XX Percent Similarity: 99.80% Conservative: 0
XX Best Local Similarity: 99.80% Mismatches: 1
XX Query Match: 99.73% Indels: 0
XX DB: Gaps: 0
XX
XX US-10-718-952-12 (1-510) x ADQ14504 (1-1533)
Qy 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyThrGluThrGluIle 20
Db 1 ATGTCATCGAGAAATTTTAAAGTAGAGAGTCTTAATGGAATGACACCGAAGCTGAGATT 60
Qy 21 GlnSerValIlyraentTyThrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTy 40
Db 61 CAGTCCGTGACAACTACGAAACACCGAAGCTTGTTCAGAGAACAGGAATGGCACTTAT 120

```

QY 41 GlnTrpIleValIleAspProIleSerValIleAsnTyrGlnPheIleThrAsnThrIleValIleAsnThrCys 60
 Db 121 GAGTGAATGTCACCAACCAATCCGCACTACCAATTAAACCAACCCAGTTGTCCA 180
 QY 61 LysLeuGlyValMetLeuValGlyTrpGlyGlyAsnAsnGlySerThrIleuThrGlyGly 80
 Db 181 AATTGGGGGTATGCTTGTGGGTTGGGGTGAACAACGGCTCTACCTCACCGGTGGT 240
 QY 81 ValIleIleAsnArgGluAspIleSerTrpAlaThrLysAspIleIleGlnIleAsn 100
 Db 241 GTTATTGCTAACAGAGGGCATTTTCATGGGCTAACAGCAAGATTCACCAACGCAAT 300
 QY 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGlu 120
 Db 301 TACTTGGCTCCCTCCACCAAGCTCAGCTATTCAGTTGAGTTCCTTCCAGGGAGGGA 360
 QY 121 IleTyrAlaProPheIleSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 Db 361 ATCTATGCCCATTCACAGAGTCTGCTTCCATGGTTAACTCGACACATTTGTTGGG 420
 QY 141 GlyTrpAspIleSerAsnMetAsnLeuValAspAlaMetAlaArgAlaLysValPheAsp 160
 Db 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCGCATGGCCAGGCAAGGTGTTGAC 480
 QY 161 IleAspLeuGlnIleGlnIleuArgProTyrTrpMetGluSerMetValProIleuProGlyIle 180
 Db 481 ATCGAATTCACAGAGCAATGAGGCTTACATGAATTCATGGTTCCATCCCGGCAATC 540
 QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluGluArgAlaAsnAsnValIleLysGly 200
 Db 541 TAGACCCGGATTTCTTCTGCTGCCACCAAGAGACCGTGCACCAACGATTAAGGCG 600
 QY 201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
 Db 601 ACAAAAGCAAGACCAAGTTGACGAATCATCAAGATCAAGGCGTTTAAAGAGGACACC 660
 QY 221 LysValAspLysValValIleuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
 Db 661 AAGGTGAGAACAGGTGTGTCTGCTGACTGCCAACACAGAGGATGATGCAATTTGGTT 720
 QY 241 ValGlyLysAsnAspThrMetGluAsnLeuLeuValAlaValAspArgAsnGluAlaGlu 260
 Db 721 GTAGGCTTAAAGACCAATGAGATCTCTTGCTGCTGCTGACAGAAATGAGGCTGAG 780
 QY 261 IleSerProSerThrIleuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
 Db 781 ATTTCTCCTTCCACCTTGTATGCAATTCCTGTGATGGAATAATGTTCTTTCATTAAAT 840
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 Db 841 GGAAGCCCTCAGAACCACTTTGTACCAAGGCGATTTGATCTGCCATCCGAGGACACT 900
 QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValIleuValAsp 320
 Db 901 TTGATTGGTGGAGATCACTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTTGGT 960
 QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
 Db 961 TTTCTTGGGGGCTGTGATCAAGCAACATCATATGATTACAAACATCATCGGGAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheAspSerLysGlyIleSerLysSer 360
 Db 1021 AATGATGGTAAATCTCGGCTCCCAACCTTCCTCCCAAGGAATAATCTCCCAAGACC 1080
 QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
 Db 1081 AAGCTTGTGAGATATGCTCAACAGCAATGCCATCTTATGAGCTGTGAACTCC 1140
 QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
 Db 1141 GACCATGTTGTATTATTAATGATATGCTTACGTAGGGGATAGCAAGAGCCATGAT 1200

QY 401 GlnTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuHisAsnThrCys 420
 Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAGAACACCATGTTTGGACACACATGT 1260
 QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
 Db 1261 GAGGATCCCTTTTGGCTGCTCTATTAATCTTGAGACTTGCTCTTCTTCTGAGCTGAGC 1320
 QY 441 ThrArgIleGlnPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
 Db 1321 ACTAGATTCAGTTTAAAGCTGAATAATGAGGAAATTCACATTCACCCAGTTGCT 1380
 QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValIleAsn 480
 Db 1381 ACCATTTCTAGATATCTGACCAAGCTCTCTGTGTTCCACCGGATACACAGTGTGAAT 1440
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
 Db 1441 GCATTGTCAAAACAGCGTCCATGCTGGAAACATATATGAGGCTTGTTGATTGGCC 1500
 QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
 Db 1501 CCAGAAATATACATGATTTCTCGAGTACAAAG 1530
 RESULT 4
 ADQ14502 standard; cDNA; 1533 BP.
 ID ADQ14502 standard; cDNA; 1533 BP.
 AC ADQ14502;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 XX inorganic phosphate; mutant.
 OS Glycine max.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "Mutant soybean myo-inositol 1-phosphate
 PN synthase #4"
 XX
 US2004128713-A1.
 PD 01-JUL-2004.
 XX
 PE 21-NOV-2003; 2003US-00718952.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-0029315.
 PR 11-MAR-2002; 2002US-0025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 XX
 DR WPI: 2004-533135/51.
 DR P-PSDB; ADQ14503.
 XX
 PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX

PS Example 8; SEQ ID NO 13; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

SO Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,986-254	Length:	1533
Score:	2624.00	Matches:	509
Percent Similarity:	99.80%	Conservative:	0
Best Local Similarity:	99.80%	Mismatches:	1
Query Match:	99.73%	Indels:	0
DB:	12	Gaps:	0

US-10-718-952-12 (1-510) x AD014502 (1-1533)

```
QY 1 MetPheIIeGluAsnPhelyValGluSerProaenValIysThrGluThrGluIle 20
DB 1 AAGTTCATGAGAATTTTAAGGAGAGAGCTCCAAATGTAAGTACCCAGACTGAGAT 60
QY 21 GlnSerValIysAsnIysGluThrGluLeuValIleGluAsnAraGsnIysThrIyr 40
DB 61 CAGTCCGCTACACACGAAACCCGAACTTTGTTCAAGAAACGAAATGGACCTTAT 120
QY 41 GlnTPRIleValIysProIysSerValAsnIyrGlnPhelyThrAsnThiIysValPro 60
DB 121 CAGTGGATTGTCAAAACCAATCCGCTACCTCACTTAAACCAACCCATGTTCCA 180
QY 61 LysLeuGluIValMetLeuValGlyIyrPglIysAsnGlySerThrIeuThrGlyGly 80
DB 181 AAATTGGGGGTATGCTTGGGTGGGTGGAAACACGCGCTTACCTCACCGGTGCT 240
QY 81 ValIleAlaAsnAraGluAspIleSerTPAlaThrIysAspIysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTAACAGAGGCGCATTTCAATGGCTACAAAGACAAAGTTCAACAGCGCAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTCCCAAGCCTCAAGCTATTCGAGTTGATCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheIysSerLeuIeuProMetValAsnProAspIleValPheGly 140
DB 361 ACTATAGCCCATCTCAAGAGTCTGCTTCAATGTGTTAATCTGACACACATGTGTGGG 420
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIysValPheAsp 160
DB 421 GGATGGGATATACGACAACTGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
QY 161 IleAspLeuGlnIysGlnLeuIyrProIyrMetGluSerMetValProIeuProGlyIle 180
DB 481 ATCGATTTCACAAGAGCTTGAAGCCTTACATGGAATTCATGGTTCATCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluGluIyrAlaAsnAsnValIleIysGly 200
DB 541 TACGACCCCGAATTCATGCTGCCAACAAAGAGAGCGTGCACAAACGTGATTAAGGGC 600
```

```
QY 201 ThrIysGlnGluIValGlnGlnIleIleIysAspIleIysValaPheIysGluAlaThr 220
DB 601 ACAAAGCAAGACAAAGTTCAAGAAATCATCAAAACATCAAGGGGTTTAAAGAACCCACC 660
QY 221 LysValaAspIysValIValLeuThrPheAlaAsnThrGluIyrGlySerAsnLeuVal 240
DB 661 AAAGTGCAAAAGGTGTGCTCTGGCTGCCCAACAGAGAGGTATAGCAATTTGGT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValaAlaValaAspAraGsnIysIleGlu 260
DB 721 GTAGGCTTAAATGACACCATGAGAAATCTTGGCTGTGTCAGAGAAATGAGGCTAG 780
QY 261 IleserProSerThrIeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTCTCTTCCACCTTGTATGCAATTCCTGTGTGATGAGAAATGTCTCTTCAATTAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyIleAspLeuAlaIleAlaAraAsnThr 300
DB 841 GGAAACCTTCAGAACACTTTGTACCAAGGCTGATGATTCCTGCGCATGCGAAGAACACT 900
QY 301 LeuIleGlyIysAspAspPheIysSerGlyIlnThrIysMetIysSerValIeuValAsp 320
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTCTACAGCAAAATGAATCTGTGTGGTAT 960
QY 321 PheIeuValGlyAlaGlyIleIysProThrSerIleValSerTyrAsnIleIysValAsn 340
DB 961 TTTCTTGGGGGCTGTATCAGCAACATCTATAGTTTCAACCATCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheAspSerIysGluIleSerIysSer 360
DB 1021 AATGATGTATGAATCTTCGCTCCACAAACCTTCGCTCCAGAAATCTCCAAAGAC 1080
QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyIysIlePro 380
DB 1081 AACGTGTGACGATATGCTACAGCAATGCCATCTCTATGAGCCTGGTGAACATCC 1140
QY 381 AspHisValIleValIleIysTyrValProTyrValGlyAspSerIysAraAlaMetAsp 400
DB 1141 GACCATGTTGTGTATTAAGTATATGTCCTTACGTAAGGGATAGCAAGAGCCATGGAT 1200
QY 401 GlnIyrThrSerGluIlePheMetGlyIysAsnThrIleValIleHisAsnThrCys 420
DB 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACACCATGTTTTCACAAACATATGT 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleLeuAlaGlyIeuSer 440
DB 1261 GAGATTCCTTTTATGCTGCTCTATTTATCTTGACCTTGCTCTTGCTGAGCTGAGC 1320
QY 441 ThrArgIleGlnPheIysAlaGluAsnGlyIysPheHisSerPheHisProValAla 460
DB 1321 ACTAGAAATCCAGTTTAAAGCTGAATAATGAGGAAATTCACATCTTCCACCCAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrIleValaProIeuValProProGlyIyrProValIleAsn 480
DB 1381 ACCATTCCTCAGTATCTTCAACCAAGGCTCTCTGTGTCCACCGGGTACCCAGTGTGTAAT 1440
QY 481 AlaLeuSerIysGlnAraAlaMetLeuGluAsnIleMetAraGlyAlaCysValaGlyLeuAla 500
DB 1441 GCATTGTCAAAAGACGCTGCAATGCTGSAAAACATTAAGAGGCTGTGTGTTGATGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluIyrIys 510
DB 1501 CCAGAGAAATTAACATGATTCCTGAGTACAAG 1530
```

RESULT 5
ADs82005 standard; cDNA, 1533 BP.

ADs82005;

18-NOV-2004 (first entry)

Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.

XX	Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid,
KW	raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX	
OS	Glycine max; cultivar Wye.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1533
FT	/*tag= a
FT	/product= "myo-inositol 1-phosphate synthase"
XX	
PN	US2003074685-A1.
XX	
PD	17-APR-2003.
XX	
PF	11-MAR-2002; 2002US-00025003.
XX	
PR	08-APR-1997; 97US-00835751.
XX	
PR	07-APR-1998; 98WO-US006822.
XX	
PA	(HITZ/) HITZ W D.
XX	(SEBA/) SEBASTIAN S A.
PI	Hitz WD, Sebastian SA;
XX	
DR	WPI; 2004-639957/62.
XX	P-PSDB; ADS82006.
XX	
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT	phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX	phytic acid and inorganic phosphate content of soybean seeds.
XX	
PS	Claim 2; SEQ ID NO 15; 34bp; English.
XX	
CC	The invention relates to an isolated nucleic acid fragment encoding a
CC	soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC	phosphate synthase having decreasing capacity for the synthesis for myo-
CC	inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC	nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC	complement, subfragment or the complement of the subfragment, operably
CC	linked to suitable regulatory sequences, where expression of the chimeric
CC	gene results in a decrease in expression of an endogenous or native gene
CC	encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC	comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC	acid content of less than 17 micromol/g, a seed content of raffinose plus
CC	stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC	greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC	from the plant, making a soybean plant with the heritable phenotype
CC	(comprising crossing LR33 or the plant comprising the chimeric gene with
CC	an elite soybean plant and selecting a progeny plant of the cross of
CC	crossing step that has a heritable phenotype as mentioned above), seeds
CC	of soybean plant made by the above method, a soy protein product derived
CC	from seeds of a soybean plant (homologous for one or more gene encoding a
CC	mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC	the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC	heritable phenotype as mentioned above), and making or producing a
CC	soybean protein product derived from seeds of a soybean plant with
CC	heritable phenotype as mentioned above. The nucleic acid is useful for
CC	altering raffinose saccharide, sucrose, phytic acid and inorganic
CC	phosphate content of soybean seeds thus leading to valuable and useful
CC	soybean products, since the presence of high concentration of raffinose
CC	oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC	when consumed by humans. The present sequence encodes a wild-type myo-
CC	inositol 1-phosphate synthase.
XX	
XX	Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX	
XX	Alignment Scores:
XX	Pred. No.: 1,98e-254 Length: 1533
XX	Score: 2624.00 Matches: 509
XX	Percent Similarity: 99.80% Conservative: 0
XX	Best Local Similarity: 99.80% Mismatches: 1
XX	Query Match: 99.73% Indels: 0

DB:	13	Gap:	0
US-10-718-952-12 (1-510) x ADS82005 (1-1533)			
QY	1 MetPheIIeGIaAsnPhelyValGIuseProaenValIyethrThrgIuThrgIuIle	20	
Db	1 ATGTTCAATCGAATAATTTAAAGTAGAGATCTTAATGTAAATGACACCGAGACTGAGATT	60	
QY	21 GInserValIYrAenThrgIuThrThrgIuLeuValIhIGIuAsnARgAnGIYThrTYr	40	
Db	61 CAGTCGGGTGTACAACTACGAACAAACCCGAACTTTGTTCAAGAAACAGGAATGGCACTAT	120	
QY	41 GInTPIIleValIyProIySserValaenrYrGInPhelyThraenThrhIseValPro	60	
Db	121 CAGTGAATGTGCAAAACCCAAATCCGTCAACTCAAAATTTAAACCAACCCACTGTTCCA	180	
QY	61 LysIeunGIyValMetLeuValGIYrTPdGIyGIyAsnAAnGIySserThrIeunThrgIyGIy	80	
Db	181 AAATTTGGGGGGATCTTGTGGGTTGGGGTGAACCAACGGCTCAACCTCAGCCGGGTGCT	240	
QY	81 ValIIealAaenARGIuAspIleSserTPAlaThrLysAspIyIleGInGInalAaen	100	
Db	241 GTTATTGCTTCAACAGAGAGGCAATTTCATGGGCTCAAAAGACAAAGATTCAACAGCCAAT	300	
QY	101 TyrPheGIySserLeuThrGInalaseValaIeArGValGIySserPheGInGIyGIu	120	
Db	301 TACTTTGGCTCCCTCACCCAAAGCTCAGCTATTGAGATTGATCTTCCAGGAGAGAA	360	
QY	121 IleyrAlaProPhelySserIeunProMetValAsnProAspAplIleValIlePheGIy	140	
Db	361 ATCTATGCCCTTCATAGAGTCTGTTCCAAAGGTTAACTTGAACAACATTTGGTTGGG	420	
QY	141 GIYrTPAspIleSserAsnMeaenLeuAlaAspAlaMerAlaARGIAlaIyValIlePheAsp	160	
Db	421 CGATGGAGATACGAACTACGTAACTGGCTGATGCCATGGCCAAAGAGGTGTTGAC	480	
QY	161 IleaPleuGIuInIySgInleuARProTYrMetGIuseMetValProleuProGIyIle	180	
Db	481 ATCGATTTGCAAGAACAGTGAAGGCTTAAATGATGATCAATGGTTCACTCCCGGAATC	540	
QY	181 TyrAspProAspPheIleAlaIaAsnGInGIuARGIAlaAsnValIleIyGIy	200	
Db	541 TACGACCCGGATTTCATTGCTGCCAACCAAGAGGCGTGCACCAACCTGATTAAAGGGC	600	
QY	201 ThrIySgInGIuInValGIuGInIleIleIySAspIleIyAlaPheIySgInAlaThr	220	
Db	601 ACAAGCAAGACCAATTCACGAATTCATCAAGACATCAAGCGTTTAAAGAGCCACC	660	
QY	221 LysValaAspIySValaIleValleuThrThralaenThrGIuARGIYrSserAsnLeuVal	240	
Db	661 AAAGTGGACAAAGGTGGTGTCTCTGTGGACTGCCAACACAGAGAGGTATAGCAATTTGGTT	720	
QY	241 ValGIyLeuAaenARThrMetGIuAsnLeuLeuAlaIaValaAspARgAnGIuAGIu	260	
Db	721 GTAGGCCCTTAATGACACCATGAGAAATCTTTGGCTGTGAGACGAATGAGCGCTGAG	780	
QY	261 IleaSserProAsrThrLeuThrAlaIleAlaCySValMetGIuAsnValProPheIleAaen	280	
Db	781 ATTTCTCTTCCACCTGTATGCCATTCCTCGTGTGATGAGAAAGTTCCTTTCATTAAAT	840	
QY	281 GlySerProGIuAenThrPheValProGIyLeuIleAspLeuAlaIleAlaARgAnThr	300	
Db	841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCGATTGATCTTGGCCATGCCGAGAACACT	900	
QY	301 LeuIleGIyGIyAspAspPheIySserGIyGInThrIySmetLysSerValleuValaAsp	320	
Db	901 TTGATGTGTGAGATGACTTCAACAGAGTGTGACAGACCAAAATCAATCTGTGTGTGAT	960	
QY	321 PheIleuValGIyAlaGIyIleIyARProThrSeriIleValSeriTYrAsnIleIeunGIyAaen	340	
Db	961 TTTCTTGTGGGGCGGTGATCAAGCCAACTTATATAGTTTACCAACCATCTGGAGAAAC	1020	
QY	341 AenAspGIyMeaenLeuSeralARProGIuThrPheArIySserIyGIuIleSeriIySser	360	

```
DB 1021 AATGATGGATGAATCTCGGCTCCACAAACCTTCGCTCCAGAAATCTCCAAAGAC 1080
QY 361 AenValValAspAspMetValAsnSerAenAlaIleuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTTGACGATATGCTCAACAGCAATGCCATCTCTTAAGAGCTGTGTAACATCCC 1140
QY 381 AspHisValValIleValIleValTyrValProTyrValGlyAspSerIleValArgLamEtAsp 400
DB 1141 GACCAATGTTGTTGTTATTAATGATATGCTTACGTAGGGGATAGCAAGAGACCATGGAT 1200
QY 401 GluTyrThrSerGluIlePheMetGlyGlyValSerThrIleValLeuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATGTTTGCACAACACATGT 1260
QY 421 GluAspSerLeuLeuAlaIleAspIleIleLeuAspLeuValIleLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTTTTACCTCTCTCTATTTATCTTGACCTTGCTCTTCTGCTAGCTGACC 1320
QY 441 ThrArgIleGlnPheValValGluValGluValGlyValPheHisSerPheHisProValAla 460
DB 1321 ACTAGATTCAGATTAAAGCTGAAATAGGGAAATTCACATTCATCCAGCTGAGTGT 1380
QY 461 ThrIleLeuSerTyrIleuThrIleValAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATTTCTCAGCTATCTGACCAAGGCTCCTCTGTTCCACCGGTACACACAGTGTGAT 1440
QY 481 AlaLeuSerIleGlnArgLamEtLeuGluValAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCAATGTCAAAACGAGGTCATCTCGGAAACATATATAGGGCTTGTGTGGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTyrIleVal 510
DB 1501 CCAGAGAAATACATGATTTCTCCAGTACAG 1530

RESULT 6
ADS82003
ID ADS82003 standard; cDNA; 1533 BP.
AC ADS82003;
XX
XX 18-NOV-2004 (first entry)
DT
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; line 29018UP03.
OS
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /tag= a
XX FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI: 2004-639957/62.
XX P-PSDB; ADS82004.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
```

```
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Example 8; SEQ ID NO 13; 34pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.
XX
XX SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,98e-254 Length: 1533
XX Score: 2624.00 Matches: 509
XX Percent Similarity: 99.80% Conservative: 0
XX Best Local Similarity: 99.80% Mismatches: 1
XX Query Match: 99.73% Indels: 0
XX DB: 13 Gaps: 0
XX
XX US-10-718-952-12 (1-510) x ADS82003 (1-1533)
XX
QY 1 MetPheIleGlnAsnPheValValGluSerProAsnValIleValTyrThrGluThrGluIle 20
DB 1 ATGTTCAATCGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
QY 21 GlnSerValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 40
DB 61 CAGTCCGTGACAACTACGAAACCAACCACTGTTTCAAGAAACGGAATGGACATAT 120
QY 41 GlnTrpIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 60
DB 121 CAGTGGATGTCAAACCAAAATCCGTCAACTACCAATTTAAACCAACCAACCAAGTTTCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyValAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATGGGGGTGATGCTTGTGGGTGGGTGGAACCAACGCTCTACCTCACCAGGTGGT 240
QY 81 ValIleAlaAsnAsnArgGluAspIleSerTrpAlaThrIleAspIleValIleGlnAlaAsn 100
DB 241 GTTATGCTAACAGAGAGGCGCATTTCAATGAGCTCAAAAGCAAGATTCAACAGCCAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTGGCTCCCTCACCACCAAGCTCAGCTATTGCAAGTTGCACTCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheValSerLeuLeuProMetValAsnProAspIleValPheGly 140
```

```

Db      |||
361 ACTATGCCCCCATTCAGAGTGTGCTTCCATGTTAATCTCGACACATTTGTTGGG 420
Qy      |||
141 GYTPAPPIESERASMETASLEUALAASPALMECALAARGALALYSVALPHEAP 160
Db      |||
421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGCAAGGTTTGGAC 480
Qy      |||
161 ILEAPLEUGLNLGAGLNLEUARGPROTYRMEGLUSEMETHALPROLEUPROGLYILE 180
Db      |||
481 ATCGATTTTCAGAGACAGTTGAGGCTTACATGGAATCCATGGGTCCACTCCCGGAATC 540
Qy      |||
181 TYRASPPIROASPHEILEALAAANGINGLUGLARGALASANAVALILEYSGLY 200
Db      |||
541 TACGACCCGGATTTTACTTCTGCGCAACAAAGAGACGCGCCAAACAGTGAATAGGGC 600
Qy      |||
201 THIRYGLINGLUGLVALINGLNIIEILEYASPIILEYVALAPHELYSGLVALATHR 220
Db      |||
601 ACMAAGCAAGACCAAGATTGACGAATCATCAAGACATCAAGGCGTTTAAAGGAGCCACC 660
Qy      |||
221 LYSVALAPLYVALVALILEUTPTTHRAAANTHGLUARGTYRSEANLEUVAL 240
Db      |||
661 AAGGTGACAGAGTGCTGTCTGTGACTGCCAACACAGAGGTATAGCAATTTGGTT 720
Qy      |||
241 VALGLYLEUASAPTHRMETGLUSENLEUALAALAVALASPARGANGLUALAGLU 260
Db      |||
721 GTAGGCTTAATGACACCATGAGAACTCTTGCGCTGCTGACAGAAATGAGCTGAG 780
Qy      |||
261 ILESERPROSETHIRLEUTYRALALEALCYVALMETGLUBANVALPROPHEILEAN 280
Db      |||
781 ATTTCTCCCTTCCACCTTGTATGCCATTGCTGTGATGGAATAATGTTCTTTCATTAA 840
Qy      |||
281 GLYSEPRROGLAASPTHIRPHEVALPROGLYLEULILEAPLEUALALEALARGAANTHR 300
Db      |||
841 GGAAGCCCTTCAAGAACCTTTGTACCAAGGCGTGAATGATCTTGCCATTCGCGAGAACCT 900
Qy      |||
301 LEULIEGLYLYASPARSPHELYSSEIRYGINTHIRYMEGLYSERVALLEUVALAP 320
Db      |||
901 TTGATTTGGTGAGATGACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGGTAT 960
Qy      |||
321 PHELEUVALGLVALAGLYILEYSPROTHRSERILEVALSERTRYANHISLEUGLYAAN 340
Db      |||
961 TTTCTTGTGGGGCTGTGATCAAGCCAACTATATGTTACCAACCATCTGGGAAAC 1020
Qy      |||
341 AENAPBGLYMERASLEUSERALAPROGINTHIRPHEARGSERLYSGLILESEIRYSSER 360
Db      |||
1021 AATGATGATATATCTCTGCGCTCCAAACCTTCGCTCCAAAGAAATCTCCAAAGAC 1080
Qy      |||
361 AENVALVALASPARSPMETVALAENSERANALALELEUTYRGLUPROGLYGLUHISPRO 380
Db      |||
1081 AACGTTGTTGACGATATGCTCAACAGCAATGCCATCTTATGAGCTGTGAAACATCCC 1140
Qy      |||
381 ASPIHSAVALVALILEYATYRVALPROTYRVALGLYASPSERLYSARGALAMEAP 400
Db      |||
1141 GACCATGTTGTTGTTATATATGATGCTTACGTAGGGGATAGCAAGACCATGAT 1200
Qy      |||
401 GLUTYRTHRSERGLILEPHEMETGLYLYVASENTHILEVALLEUHIASANTHRCYG 420
Db      |||
1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATGTTTGCACAACACATGT 1260
Qy      |||
421 GIUAPSPSERLEUVALAALAPROILEILELEUAPLEUVALLEUVALAGLUSEUSER 440
Db      |||
1261 GAGGATTCCTTTTACTGCTCTCTATATCTTGACATGCTGCTCTTCTTGTCTAGCTGACC 1320
Qy      |||
441 THIRARGILEGINPHELYSVALAGLUBANGLUGLYLVSPHENISSEPHENIASPROVALA 460
Db      |||
1321 ACTAGAAATCAGTTTAAAGCTGAATAATGAGGAAATTCACATTCACACCAAGTGTGT 1380
Qy      |||
461 THIRILELESETTYRLEUTHIRLYSVALAPROLEUVALPROPROGLYTHIRPROVALVALAN 480
Db      |||
1381 ACCATTCTCAAGTATCTGACCAAGGCTCTCTGTTTCCACCGGATACACACATGTGTGAT 1440
Qy      |||
481 ALALUSERLYSGLARGALAMELEUGLUBANILEMERARGALCYVALGLYLEUVALA 500

```

```

Db      1441 GCATTGTCAAGACAGCTGCAATGCTGGAATAATGAGGCTGTGTGATTGCC 1500
Qy      501 ProGUaenAenMeTlIeUeGluTyTlyS 510
Db      1501 CCAGAAATATACATGATTTCTCGATGACAAAG 1530
RESULT 7
ADQ14498
ID ADQ14498 standard, cDNA, 1533 BP.
XX
AC ADQ14498;
XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.
XX
KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
OS Glycine max.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /*tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #2"
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006832.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
DR WPI: 2004-533135/51.
DR P-PSDB; ADQ14499.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Example 8, SEQ ID NO 9, 48bp; English.
XX
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful

```

CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,65e-251	Length:	1533
Score:	2593.00	Matches:	502
Percent Similarity:	99.22%	Conservative:	4
Best Local Similarity:	98.43%	Mismatches:	4
Query Match:	98.56%	Indels:	0
DE:	12	Gaps:	0

US-10-718-952-12 (1-510) x ADQ14498 (1-1533)

```

QY 1 MetPheileGluAsnPhelysValGluSerProAsnValysrThrGluThrGluile 20
DB 1 ATGTCATCGAATAATTTTAAAGTTGAGTGTCTTAATGCAATACCAAGACTAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCTGTACAACTACGAAACCAACGAACTTGTTACAGAAACAGAAATGGCACTAT 120
QY 41 GlnTPPileValLysProLysSerValAsnTyrGlnPheLysThrAsnThiSValPro 60
DB 121 CAGTGGATTGCAAAACCAAAATCTGCAAAATGCAAAATTTAAACCAACATTCATGTTCT 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTAAAGGGTAATGCTTGTGGGTGGGTGGAAACAAAGCTCAACCTCCAGGTGCT 240
QY 81 ValIleAlaAsnArgGluAspIleSerTrrAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTTAACCGAGAGGGCATTTTCATAGGCTACAAAGACAAAGATTCAACAACTCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTCAACCAAGCTTCAGCTATCCAGTGGGTGCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCCCATTCAGAGCTGCTCTCCATGCTTAACCTGACGACATGTGTGGG 420
QY 141 GlyTPPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCATGGCCAGAGGCAAAAGGTTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTGTGCAAGACAGTTGAGGCTTTCATGAAATTCATGCTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluGluArgAlaAsnAsnValIleLysGly 200
DB 541 TATGACCCCGAATTTCTATCTCTCCAAACAAAGAGCGGCCAAACATCTCATCAAGGCC 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAAAGCAAGACCAAGTTCAACAATCATCAAAAGCATCAAGCGCTTAAAGAAACCAAC 660
QY 221 LysValAspLysValValValLeuTrrPThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGGACAAAGGTGGTGTACTGTGAGCTGCAACACAGAGAGTACAGTAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTAAATGACCACTGAGAAATCTCTTGCTGCTGTGAGCAAGAAATAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTGCTCTTCCACCTGTATGCAATGCTGTGATTATGAAATAATGTTCTTTCAATTAAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300

```

```

DB 841 GGAAACCCCTCGAAACACTTTTGTACCAAGGCTGATTTGATCTTGGCATCGGAGAACACT 900
QY 301 LeuileGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGAAGATGACTTCAGAGTGGTGCAGACCAAAATGAATATCTGTGTGGTTCAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTCTTGGGGGGCGTGGTATMACCAACATCTATGCACTTTCATCAACCACTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGGTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGGAATTCACAGAGC 1080
QY 361 AsnValValaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCGCTGGTGAACATCA 1140
QY 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTGTTATTAAGTATGCTTACCTAGGGGACAGCAAGAGCCATGGAT 1200
QY 401 GluTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTTATGAGGTGGAGAAAGACACCAATGTTTTCACAAACATGC 1260
QY 421 GluAspSerLeuLeuAlaIleProIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGATTCCCTCTTAGCGCTCTTATATCTTGAACCTTGGCTTCTTGGCTGAGCTGAGC 1320
QY 441 ThrArgIleGlnPheLysValGluAsnGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTGAATCGAGTTTAAAGCTGAAATGAGGAAATTCACCTATCCACCCAGTGTCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTrrProValValAsn 480
DB 1381 ACCATCTCAGCTACCTACCAAGGCTCTCTGTTCCACGGGTACCAAGTGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTGTCAAAAGACGCGCAATGCTGGAATAAATGAGCGCTTGTGTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1501 CCAGGATTAACATGATTCCTGAGTACAAAG 1530

```

RESULT 8
ID ADS81999 standard; cDNA; 1533 BP.
XX
AC ADS81999;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29004JDP01.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
FN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PE 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX
PI Hitz WD, Sebastian SA;
XX
XX WPI: 2004-639957/62.
DR P-PSDB; ADS82000.
XX
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
DR
DR
XX
XX Example 8; SEQ ID NO 9, 34p; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.
CC
CC
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,65e-251 Length: 1533
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.56% Indels: 0
DB: 13 Gaps: 0
US-10-718-952-12 (1-510) x ADS81999 (1-1533)
QY 1 MetPheIleGluAspPheIysValIgluSerProAsnValIleuTyThrGluThrGluIle 20
Db 1 ATGTCATCGAGAAATTTAAAGGTGAGTCTCCAAATGTAAGTACCCGAGACTGAGATT 60
QY 21 GlnSerValIyTrAsnTyrgIuThrGluLeuValIleGluAsnAryAsnGlyThrTy 40
Db 61 CAGTCCGTTGACACACGAAACCCGAACTTGTTCAAGAGACGAGATGGCACTAT 120
QY 41 GlnTTPilIeValIyPProIySerValIAsnTyrgIuPheIySerThrAsnThiIValIPro 60
Db 121 CAGTGGATGTCGAAACCCAAATCTGTCAAATGCAATTTAAACCAACATTCATGTTCT 180
QY 61 LysIleGluIValMetIeuValIgluTyrgIyValIAsnAryGlySerThrIeuThrGlyGly 80

Db 181 AATATTAGGGGTAAATGCTTGTTGGGGTGGGAAACAAGCTCAACCTCAACGGTGGT 240
QY 81 ValIleAlaAsnAryGluAspIleSerTrpAlaThrIyAspIyIleGlnIleAlaAsn 100
Db 241 GTTATTGGCTAACCGAGAGGCGCATTTTCATGGGCTTCAAAAGCAACAAATTCACAAACGCAAT 300
QY 101 TyrPheGlySerIleuThrGluAlaSerIleAlaArgValIgluSerPheGlnIyGluIu 120
Db 301 TACTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGCTCTTCCAGGGAGAGGAA 360
QY 121 IleTyAlaProPheIySerIeuIeuProMetValIAsnProAspAspIleValPheGly 140
Db 361 ATCTATGCCCATTTCAAGAGCGCTGCTTCCAATGTTAACTCGAGACATTTGTTGGG 420
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIyValIlePheAsp 160
Db 421 CGATGGGATATCAGCAATCATGAACTGCTGATGCAAGCGCAAGGCAAGGCTGTTGAC 480
QY 161 IleAspIleGlnIyGluLeuAryProTyTrMetGluSerMetValIProIeuProGlyIle 180
Db 481 ATCGATTGCAAGACAGTGGAGGCTTACATGAAATCATGCTTCCACTCCCGGAAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnIyGluAryAlaAsnAryValIleIyGly 200
Db 541 TATGACCGGATTTCAATGCTGCTCCCAACAGAGAGCGGCCCAACAAGCTCATCAAGGCG 600
QY 201 ThrIyGlnIyGluIValIgluGlnIleIleIyAspIleIyValIlePheIyGluAlaThr 220
Db 601 ACAAAAGCAAGAGCAAGTTCACAAATTCATCAAAAGCATCAAGGCGCTTAAAGCAACCCACC 660
QY 221 LysValAspIyValIValIleuThrIleAlaAsnThrGluArySerAsnIeuVal 240
Db 661 AAAGTGCACAAAGTGTGTTACTGTGACCTGCAACAGAGAGTACAGTAATTTGGTT 720
QY 241 ValIyGlyLeuAsnAspThrMetGluAsnIeuAlaAlaValAspAryAsnIuIaglu 260
Db 721 GTGGGCTTAAATGACCATGAGATCTTTGGCTGCTGTGAGACGAAATAGGCTAG 780
QY 261 IleSerProSerThrIeuTyAlaIleAlaCysValMetGluAsnValIProPheIleAsn 280
Db 781 ATTTCTCTTCACCTTGATGCCATTCCTGTGTTATGAAAGTCTCTTCAATTAAT 840
QY 281 GlySerProGlnAsnThrPheValIProGlyIleuIleAspIleAlaIleArgAsnThr 300
Db 841 GGAAACCCCTCAGAAACATTTGTACAGGCGCTGATGATCTTGGCATCGCGAGAACAT 900
QY 301 LeuIleGlyIyAspAspPheIySerGlyIuThrIyMetIySerValIleuValAsp 320
Db 901 TTGATTGGTGAATGACTTTCAGAGTGTCTCAGACCAAAATGAATCTGTGTGGTTGAT 960
QY 321 PheIeuValIgluAlaGlyIleIyAspProThrSerIleValIleSerTyAsnIleuGlyAsn 340
Db 961 TTCCTTGGGGGGCGTATATCAAGCCAACTATATAGTACATGACATCGATGGGAAAC 1020
QY 341 AsnAspGlyMetAsnIeuSerAlaProGlnThrPheIySerIyGluIleSerIySer 360
Db 1021 AATGATGTATGAATCTTGGCTCCACAAATTTCCGTTCCAGAAATCTCCAAAGAGC 1080
QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleuTyrgIuProGlyIyIuIAspPro 380
Db 1081 AACGTTGTGATGATATGCTCAACGCAATGCACTCTATGACCTGGTGAACATCCA 1140
QY 381 AspIleValIleValIleIySerTyValIProTyValIgluAspSerIyAspAlaMetAsp 400
Db 1141 GACCATGTTGTTGTTAATTAAGTATGCTTACGTAAGGAGACAGCAAGAGCCATGAT 1200
QY 401 GlyTyThrSerGluIlePheMetGlyIyIyAsnThrIleValIleuIleAsnThrCys 420
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGACACATTTGTTTSCAACAAACATGC 1260
QY 421 GluAspSerIleuIleAlaIProIleIleuAspIleuValIleuIleAlaGlyIuIeuSer 440
Db 1261 GAGGATTCCTCTTACGCTGCTCTATATCTTGAATCTTGCTTCTGCTGAGCTCAGC 1320

QY 441 ThrArgIleGlnPheIysAlaGluasnGluGlyIysPheHisSerPheHisProValAla 460
DB 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTCATCTCATTCACCCAGTTGCT 1380
QY 461 ThrIleuSerTyrIleuThrIysAlaProIleuValProProGlyThrProValAlaasn 480
DB 1381 ACCATCCCTCAGCTACCTCCACCAAGGCTCCTCGGTGCCACCGGTACACCACTGGTGAAT 1440
QY 481 AlaIeuSerIysGlnArgAlaMetIleuGluAsnIleMetArgAlaCysValGlyIleuAla 500
DB 1441 GCAATGTCAAGCAGGCTGCAATGCTGGAAGAACTATATAGAGGCTTGTTGGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleuGluIlyrIlys 510
DB 1501 CCAGAGAAATACATGATCTCTCGAGTACAG 1530
RESULT 9
ADS81993
ID ADS81993 standard; cDNA, 1533 BP.
XX
AC ADS81993;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
XX P-PSDB; ADS81994.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 2; SEQ ID NO 1; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds

CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homologous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.
XX
SQ Sequence 1533 BP, 429 A, 344 C, 368 G, 392 T, 0 U, 0 Other;
Alignment Scores:
Pred. No.: 2,65e-251 Length: 1533
Score: 2593.00 Matches: 502
Percent Similarity: 99.32% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.56% Indels: 0
DB: Gaps: 0
US-10-718-952-12 (1-510) x ADS81993 (1-1533)
QY 1 MetPheIleGluAsnPheIysAlaGluSerProAsnValIlyrIlyrThrGluThrGluIle 20
DB 1 ATGTTTCATCGAAGATTTTTAAAGTTGAGTGTCTATGTGAAGTACACCGAATCGAAGATT 60
QY 21 GlnSerValIlyrAsnTyrGluThrThrGluIleuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCTGACACTACGAAACCCGAACTGTTCACGGAACGGAATGCACTAT 120
QY 41 GlnTrpIleValIlyrProIysSerValAsnTyrGlnPheIysTrpAsnThrHisValPro 60
DB 121 CAGTGGATTGTCAAAACCCAAATCTGTCAATATGCAATTTAAACCAATCATCTTCT 180
QY 61 LysIleuGlyValMetIleuValGlyTyrGlyGlyAsnAsnGlySerThrIleuThrGlyGly 80
DB 181 AAATTAGGGGTAAATGCTTGTGGGTGGGTAACCAACGCTCAACCTCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrIysAspIlyrIleGlnAlaAsn 100
DB 241 GTTATGCTTAACCGAGAGGCAATTCATGAGCTCAAAAGACAAAGATTCAACAGCCAAAT 300
QY 101 TyrPheGlySerIleuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTCACCAAGCGCTCAGCTATCCGAGTGGGTCTTCCAGGAGAGAA 360
QY 121 IleTyrAlaProPheIysSerIleuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTCACAGAGCTGCTTCCAAATGTTAACTCTGACGACATGTGTTGGG 420
QY 141 GlyTrpAspIleSerAsnMetAsnIleuAlaAspAlaMetAlaArgAlaIlyrValPheAsp 160
DB 421 GGAATGGATATCAGCAATATACCTGCTATGCTATGCTAGCCAGGCAAGAGTTTATAC 480
QY 161 IleAspIleuGlnIysGlnIleuArgProTyrMetGluSerMetValProIleuProGlyIle 180
DB 481 ATCATGTTGCAAGAGCAAGTTAGAGGCTTACATGATATCAAGCTTCCATCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluIlyrGluAlaAsnAsnValIleIysGly 200
DB 541 TATGACCCGGAATTCATGCTGCAACCAAGAGAGGCTGCCAACAAGTATCAAGAGGC 600
QY 201 ThrIysGlnGluGlnValGlnGlnIleIleIysAspIleIysAlaPheIysGluAlaThr 220
DB 601 ACAAAGCAAGCAAGTTCAACAAATCATCAAGACATCAAGAGCGGTTTAAAGAAAGCCACC 660

QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
 DB 114 CAGTCGCTGACAACTACCAACCAACGAACTGTGTACAGAAACAGAAATGGCACTTAT 173
 QY 41 GlnTPILeValLysProLysSerValAsnTyrGlnPheLysThrAsnThrHisValPro 60
 DB 174 CAGTGGATTGTCAAAACCAAAATGTGCAAAATGCAAAATTTAAACCAAACTCCATGTCTCT 233
 QY 61 LysLeuGlyValMetLeuValGlyTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
 DB 234 AAATTAGGGGTATATGTTGTGGGTTGGGTTGGAAACAAGCGCTCAACCTTCACCGGTGT 293
 QY 81 Val11Lea1aAsnArgGluAsp11LeuSerTPA1aThrLysAspLys11GlnGln1aAsn 100
 DB 294 GTTATTGTCTTACCGAGAGGCAATTCATAGGCTACAAAGACAAGATTCACCAAGCAAT 353
 QY 101 TyrPheGlySerLeuThrGln1aSer1a1LeaGlyValGlySerPheGln1aGlyGlu 120
 DB 354 TACCTTGGCTCCCTCAACCAAGCTTCAGCTATCCAGTTGGGTCTTCCAGGAGAGGAA 413
 QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAsp11Lea1aPheGly 140
 DB 414 ATCTATGCCCCCATTCAGAGCGCTGCTTCCAAATGTTTAAACCTGACGACATTTGTGTTGG 473
 QY 141 GlyTPAsp11LeuSerAsnMetAsnLeu1aAsp1a1aMetAlaArg1aLysValPheAsp 160
 DB 474 GGATGGGATATACGAACATGAACCTGCTGATGTCATGGCCAGGCAAAAGGTGTTGAC 533
 QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGly1Le 180
 DB 534 ATCGATTTCAGAAAGAGTTGAGGCTTACATGAATTCATCTTCACCTCCCGGAATC 593
 QY 181 TyrAspProAspPhe11Lea1a1aAsnGlnGluArg1aAsnAsnVal11LeuGly 200
 DB 594 TATGACCCGGATTTCATCTCTCCCAACCAAGAGGCGTCCAAACACCTCATTCAGAGGC 653
 QY 201 ThrLysGlnGluGlnValGlnGln11LeuLysAsp11LeuValPheLysGlu1aThr 220
 DB 654 ACAAAAGCAAGACCAAGTTCACAAATCATCAAAAGCATCAAGCGCTTAAAGAAACCAAC 713
 QY 221 LysValAspLysValValLysLeuTPHra1aAsnThrGluArgTyrSerAsnLeuVal 240
 DB 714 AAGAGTGACAAAGGTGTGTACTGTGAGCTGCACACAGAGAGTACAGTAATTTGGTT 773
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuVal1a1aValAspArgAsnGlu1aGlu 260
 DB 774 GTGGGCTTAATGACACCATGAGAAATCTCTTGCGCTGTGTGACAGAAATAGCGTGA 833
 QY 261 IleSerProSerThrLeuTyrAla11Lea1aGlyValMetGluAsnValProPhe11Lea 280
 DB 834 ATTTCTCTCCATCCACTTGTATGCAATGCTGTGTATGAAATGTTCTCTTCAATTAT 893
 QY 281 GlySerProGlnAsnThrPheValProGlyLeu11LeaPhe1a1a11Lea1aArgAsnThr 300
 DB 894 GGAAGCCCTCAACAACCTTTGTACCAAGGCGTGAATTCCTTGCATCGCGAGAACCT 953
 QY 301 Leu11LeuGlyLysAspAspPheLysSerGly1aThrLysMetLysSerValLeuValAsp 320
 DB 954 TTGATTGGTGGAGATGACTTCAGAGTGTGACAGCAAAATGAAATCTGTGTGGTTGAT 1013
 QY 321 PheLeuValGlyAlaGly11LeuLysProThrSer11Lea1aSerTyrAsnHisLeuGlyAsn 340
 DB 1014 TTCTTGTGGGGGTGGTGTCAAGCAACATCTATATAGTACATTAACACCATCTGGGAAAC 1073
 QY 341 AsnAspGlyMetAsnLeuSer1aProGlnThrPheArgSerLysGlu11LeuSerLysSer 360
 DB 1074 AATGATGTGATGAATCTTGGGCTCCACAACTTCCGTTCCAGAGAAATCTCCAAAGGC 1133
 QY 361 AsnVal1a1aAspAspMetValAsnSerAsnAla11LeuTyrGluProGlyGluHisPro 380
 DB 1134 AACGTTGTGATGATAGTGTCAACAGCAATGCAATCTCTTATGACCTGGGTGACATCA 1193
 QY 381 AspHisValVal1a11LeuLysTyrValProTyrValGlyAspSerLysArg1aMetAsp 400

DB 1194 GACCATGTTGTGTATTAATGATGTGCTTACATGGGGAACACCAAGACCATGAT 1253
 QY 401 GluTyrThrSerGlu11LeaPheMetGlyGlyLysAsnThr11LeuHisAsnThrCys 420
 DB 1254 GAGTACATTCAGAGATTTTATGAGTGTGGAGTGGAAAGACCAATGTTTTCACAAACATATC 1313
 QY 421 GluAspSerLeuLeu1a1aPro11LeuAspLeuVal1LeuLeu1aGlyLeuSer 440
 DB 1314 GAGGATTCCTCTTATAGCTGCTCTTATATCTTGGACTTGCTGCTTCTTGTGAGCTCAGC 1373
 QY 441 ThrArg11GlnPheLysValGluAsnGlyGlyLysPheHisSerPheHisProVal1a 460
 DB 1374 ACTGAATCGAGTTTAAAGCTGAATAATGAGGAAATTCACATTCACACCCAGTTGCT 1433
 QY 461 Thr11LeuSerTyrLeuThrThrLysAlaProLeuValProProGlyThrProVal1aAsn 480
 DB 1434 ACCATCTCAGCTACCTACCAAGGCTCTCTGTGTTCCACCGGGTACCAAGTGTGAT 1493
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsn11MetArgAlaCysValGlyLeu1a 500
 DB 1494 GCATTGTCAAAAGACCGTGCATGCTGGAATAATGAGGGCTTGTGTGATTTGGCC 1553
 QY 501 ProGluAsnAsnMet11LeuGluTyrLys 510
 DB 1554 CCAGAGAAATTAATGATTCGAGTACAAAG 1583
 RESULT 11
 AAV62440
 ID AAV62440 standard; cDNA, 1782 BP.
 XX
 AC AAV62440;
 XX
 DT 17-OCT-2003 (revised)
 DT 02-FEB-1999 (first entry)
 XX
 DE Soybean wild-type myo-inositol 1-phosphate synthase cDNA.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
 KW phytic acid; ds.
 XX
 OS Glycine max; line LR13.
 XX
 FH Key Location/Qualifiers
 FT CDS 54..1586
 FT /*tag= a
 XX
 FN W09845448-A1.
 XX
 PD 15-OCT-1998.
 XX
 PF 07-APR-1998; 98MO-US006822.
 PR 08-APR-1997; 97US-00835751.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 PI Hitc WD, Sebastian SA;
 DR WPI; 1998-568353/48.
 DR P-PSDB; AAW79740.
 XX
 PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
 PT for generating plants with altered levels of e.g. raffinose, stachyose,
 PT phytic acid, etc.
 PS
 PS Example 5; Page 44-45; 63pp; English.
 CC This is the nucleotide sequence of cDNA encoding the wild-type soybean
 CC myo-inositol 1-phosphate synthase (MI 1-PS) present in clone psbmi-1ps
 CC (ATCC 97970). The clone was isolated from a cDNA library of soybean line
 CC LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
 CC thaliana. MI 1-PS is involved in glucose metabolism to phytic acid.

PD 15-OCT-1998.
XX
XX 07-APR-1998; 98WO-US006822.
XX
XX 08-APR-1997; 97US-00835751.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI, 1998-568353/48.
XX P-PSDB; AAM79741.
XX
XX Soybean plants containing altered myo-inositol-1-phosphate gene - useful
XX for generating plants with altered levels of e.g. raffinose, stachyose,
XX phytic acid, etc.
XX
XX Example 5; Page 48-49; 63pp; English.
XX
XX This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
XX inositol-1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
XX clone was isolated from a cDNA library of soybean line LR33 by PCR
XX amplification (see AAV62441-42). Line LR33 was obtained by chemical
XX mutagenesis of wild-type soybean genome and as a reduced raffinose
XX saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
XX acid, raffinose and stachyose. Sequencing revealed a single base change
XX mutation (G to T at base 1241) in the LR33 sequence when compared to the
XX wild-type sequence (see AAV62440). The mutation results in a seed
XX phenotype of very low raffinose saccharide sugars, very high sucrose and
XX low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
XX raffinose saccharide, sucrose, phytic acid and inorganic phosphate
XX content of soybean seeds, leading to useful soybean products, e.g. a seed
XX phytic acid content of less than 14.5 ug/g, a seed content of raffinose and
XX stachyose combined of less than 14.5 ug/g, and a seed sucrose content
XX greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 8 476-251 Length: 1533
XX Score: 2588.00 Matches: 501
XX Percent Similarity: 99.02% Conservative: 4
XX Best Local Similarity: 98.24% Mismatches: 5
XX Query Match: 98.37% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-10-718-952-12 (1-510) x AAV62443 (1-1533)
XX
XX QY 1 MetPheileGluAsnPhelyValGluSerProAsnValLysTyrThrGluThrGlu 20
XX DB 1 ATGTTCAATCGAAGATTTTAAAGGTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
XX QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
XX DB 61 CAGTCCGTTGATACACTAGCAACACCGAACTGTTCAGAGAAACGAAATGGCACTAT 120
XX QY 41 GlnTrrpileValLysProLysSerValAsnTyrGlnPhelyThrAsnThrHisValPro 60
XX DB 121 CAGTGGATTGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATCTCAATTTCTT 180
XX QY 61 LysLeuGlyValMetLeuValGlyTrrpGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
XX DB 181 AAATTAGGGGTAAATGCTTGTTGGGTGGGTGGAACAAAGCTCAACCTGACCGGTGT 240
XX QY 81 ValIleAlaAsnArgGluAspIleSerTrrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
XX DB 241 GTTATTGCTAAACCGAGAGGCAATTCATGAGCTACAAAGACAAATTCAAACCAAT 300
XX QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlnGlu 120
XX DB 301 TACTTTGGTCTCCCTCAACCAAGCTTCAGCTATCCAGTTGGGTCTCTTCCAGGAGAGGAA 360
XX QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140

DB 361 ATCTATGCCCATTCAGAGAGCTGCTTCCAAATGGTTAACTCTGACGACATGTGTTGGG 420
QY 141 GYTrrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGATATACCAACATCAATCACTGCTGATGCCATGGCCAGAGGAGAGTGTTCAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTGGCAGAGAGCTTGAGGCTTACATGGAATTCATGATCTCCACCTCCCGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TATACCCGATTTTCATTGCTGCGCAACCAAGAGAGGTGCAACAGTATCAAGGAC 600
QY 201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAGCAAGAGCAAGTTCACCAATCATCAAAACATCAAGGCGTTTAAAGAAAGCCACC 660
QY 221 LysValAspLysValValLeuTrrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGCACAGTGGTGTGACTGTGACCTGCCAACACAGAGAGTACAGTAATTTGGTT 720
QY 241 ValGlyLeuAsnAspTrrpMetGluAsnLeuLeuAlaValaAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTAAATGACACCATGAGAAATCTTGGACTGTGGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerTrrpLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCACTTGTATGATGATGCCATTCCTGTGTATGAGAAATGTTCTTCATTAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGGCTGATGATGATCTTGCCATGCGAGAGAACT 900
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGAATGATGACTTCACAGATGCTCAGACCAAAATGAAATCGTGTGGTGTAT 960
QY 321 PheLeuValGlyValaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTCTTGGGGGGCTGGATATGACCAACATCTATAGTCACTTACACCACTGGGAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGGTATGAATCTTTCGGCTCCAAACTTTCCTCCAGGAATCTCCAAAGAGC 1080
QY 361 AsnValValaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTGATGATATGATGACAAAGCAATGCAATCTCTATGAGCCTGGTGAACATCCA 1140
QY 381 AspHisValValaIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCAATGTTGTTATTAAGTATATGCTCTTAAGTGGGACACAGAAATAGGCCATGAT 1200
QY 401 GluTyrTrpSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAAGACCAATGTTGTTTSCAACACAGATGC 1260
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTAAGCTCTATATATCTTGAACCTTGCTCTTCTTGCTGAGCTGAGC 1320
QY 441 ThrArgIleGlnPhelyValaGluAsnGlnGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGAAATCGAGTTTAAAGCTGAAATATGAGAAATTCATCTCAATCCACCAAGTCTCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATCTCAGACTACCTACCAAGGCTCTGTGTTCCACGGGTACACAGTGGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500

DB 1441 GCATTGTCAGACGCGTCGAAATGCTGAAACATATATAGAGGCTTGTTGGATTGGCC 1500
QY 501 ProGUAsnAsnMetIleuGluTyrTyr 510
DB 1501 CCAGAGAAATACATGATCTCTCAGATCAAG 1530
RESULT 13
ADQ14494
ID ADQ14494 standard; cDNA; 1533 BP.
XX ADQ14494;
XX ADQ14494;
XX ADQ14494;
XX 23-SEP-2004 (first entry)
XX
XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.
DE
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
XX Glycine max.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 1.1533
FT CDS
FT /*tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #1"
FT replace(1241,G)
FT mutation
FT /*tag= b
PN
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00293315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX MPI: 2004-533135/51.
DR P-PSDB; ADQ14495.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Claim 10; SEQ ID NO 5; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with

CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8,47e-251 Length: 1533
Score: 2588.00 Matches: 501
Percent Similarity: 99.02% Conservative: 4
Best Local Similarity: 98.24% Mismatches: 5
Query Match: 98.37% Indels: 0
DB: 12 Gaps: 0
US-10-718-952-12 (1-510) x ADQ14494 (1-1533)
QY 1 MetPheIleGluAsnPheIleValGluSerProAsnValIleTyrThrGluThrGluIle 20
DB 1 ATGTTTCATCGAATTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluValIleGluAsnArgAsnGlyTyrTyr 40
DB 61 CAGTCGCTGATACACTACGAAACCAACCACTGTTCAAGAGAACAGAACTGACCTAT 120
QY 41 GlnTrrIleValIleValProIleSerValAsnTyrGlnPheIleThrAsnThrIleValPro 60
DB 121 CAGTGATTTGTCAAACCCAAATCTGTCAAAATGCAATTTAAACCAATTCATGTTCTT 180
QY 61 LysLeuGlyValMetLeuValGlyTrrPglyAsnAsnGlySerThrIleThrGlyGly 80
DB 181 AAATTAGGGGTAAATCTTGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGT 240
QY 81 ValIleAlaAsnArgGluAsnIleSerTrrPalaThrIleValIleGlnIleAlaAsn 100
DB 241 GTTATTGCTAACCGAGAGGCGATTTCATGGGCTTCAAAAGAACAAATTCACCAAGCCT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleAlaGlySerPheGlnGlyGlu 120
DB 301 TACTTGGCTCCCTCACCACCAAGCTTCACTGATCCAGTGGTCCCTCCAGGAGAGGAA 360
QY 121 IleTrrAlaProPheIleSerLeuIleProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTTAAAGAGCTGCTCTCAATGTTAACTTCAGACCAATTTGTTGGG 420
QY 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
DB 421 GGATGGATATCAGCAATGAACTGCTGATGCCATGGCCAGGCAAGGTTGTTGAC 480
QY 161 IleAspLeuGlnIleValLeuArgProTrrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTTGCAAGACAGCTGAGGCTTACATGGAATCATGCTTCCACTCCCGGAATTC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluGlyAlaAsnValIleIleGly 200
DB 541 TATGACCCGGAATTTATGCTGCAACCAAGAGGCGGTGCAACCAAGCTTCAAGGGC 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluIleThr 220
DB 601 ACAAAAGCAAGAGCAAGTTCAACAATATCAAAAGCATCAAGGCTTTAAGAAACCAAC 660
QY 221 LysValAspLysValValIleValLeuTrrPheAlaAsnThrGluArgTrrSerAsnLeuVal 240
DB 661 AAAGTGAACAAGTGTGTGTGCTGAGCTCCCAACACAGAGAGGTACAGTAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluIleGlu 260
DB 721 GTGGGCTTAAAGACACCATGAGATCTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCCTCTTCAACCTTGTATGCACTTGTGTGTATGAGAAATGTTCTTTATTAAT 840

QY	281	GlySerP	ProGlnAsnThrPheValP	ProGlyLeuIleAspLeuAlaIleAlaArgAsnThr	300
Db	841	GGAAAGCCCTCAGAAACACTTTTGTACACAGGGCGTGAATGATCTTGGCATGCCAGGAAACACT			900
QY	301	LeuIleGlyValAspAspPheLeuSerGlyGlnThrIlySmcLysSerValIleuValAsp			320
Db	901	TTGATTTGTTGAGATGACTTTCAGAGTGGTCCAGACCAAAAGAAATCTGTGTGGTTGAT			960
QY	321	PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTYZAsnHileuGlyAsn			340
Db	961	TTCCCTGTGGGGGGCGGTATCAGCCAAACATCTATAGTCACATTACACCATCTGGGAAAC			1020
QY	341	AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer			360
Db	1021	AATGATGGTAAGAATCTTTCGGCTCCAAAACCTTCCGTTCCAGGAATCTCCAAAGGC			1080
QY	361	AsnValValaPAPAspMetValAsnSerAsnAlaIleLeuTYRglnUpProGlyGlnHileu			380
Db	1081	AACGTTGTGATGATATGATGTCACACGACATGCCATCTCTATAGACCTGGTGAACATCCA			1140
QY	381	AspHisValValIleLysTYRValProTYRValGlyAspSerLysArgAlaMetAsp			400
Db	1141	GACCATGTGTTGTTGTTATTAAGATGTGCTTACGTAGGGGACAGCAATAGAGCCATGGAT			1200
QY	401	GluTYRThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys			420
Db	1201	GAGTCACTTCAAGATATTCATGCGGTGGAAGAGACCATTTGTTTGCAACACACATGC			1260
QY	421	GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuLeuAlaGlnLeuSer			440
Db	1261	GAGGATTCCTCTTAGCTGCTCTCTATATACCTTGACCTTGCTCTTCTTGCTGAGCTCAGC			1320
QY	441	ThrArgIleGlnPheLysAlaGluAsnGlnGlyLysPheHisSerPheHisProValAla			460
Db	1321	ACTAGAAATCAGACTTTAAAGCTGGAATAATAGGGAAATTCCTCATCTCCACCAGTTGCT			1380
QY	461	ThrIleLeuSerTYRLeuThrIlyAlaProLeuValProProGlyTYRProValValAsn			480
Db	1381	ACCATCCCTCAGCTACCTCACCACAGGCTCTCTCGTTCCACCGGGTACACACAGTGGTGAAT			1440
QY	481	AlaIleuSerLysGlnArgAlaMetLeuGlnAsnIleMetAlaGlnAlaCysValGlyLeuAla			500
Db	1441	GCATTGTCAAAAGCACACGTCGCAATCTGGAAAAACATTAATGAGGGCTGTGGTGGATTGCC			1500
QY	501	ProGluAsnAsnMetIleLeuGlyTYRLeuLys			510
Db	1501	CCAGAGAAATACATGATTTCTCGAGTACAAAG			1530
RESULT 14					
ADSD81997	ID	ADSD81997	standard; cDNA; 1533 BP.		
XX	AC	ADSD81997;			
XX	XX	18-NOV-2004 (first entry)			
XX	DE	Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.			
XX	KM	Soybean; plant; myo-inositol 1-phosphate synthase; ssi gene; phytic acid;			
XX	XX	raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.			
OS	OS	Glycine max; line LR33.			
XX	XX	Synthetic.			
FT	FT	Key	Location/Qualifiers		
FT	FT	CDS	1..1533		
FT	FT	/tag=	a		
FT	FT	/product=	"myo-inositol 1-phosphate synthase"		
FT	FT	replace(1188,G)			
XX	XX	/tag=	b		
PN	US2003074685-A1.				

D		17-APR-2003.			
XX	P	11-MAR-2002; 2002US-00025003.			
XX	F				
PR	O	08-APR-1997; 97US-00835751.			
PR	A	07-APR-1998; 98MO-US006822.			
XX					
PA	(HITZ/) HITZ W D.				
PA	(SEBA/) SEBASTIAN S A.				
XX					
PI	Hitz WD, Sebastian SA;				
XX					
DR	WPI; 2004-639957/62.				
DR	P-PSDB; ADS81998.				
XX					
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-				
PT	phosphate synthase, useful for altering raffinose saccharide, sucrose,				
ET	phytic acid and inorganic phosphate content of soybean seeds.				
PS	Claim 8; SEQ ID NO 5; 34pd; English.				
XX					
CC	The invention relates to an isolated nucleic acid fragment encoding a				
CC	soybean myo-inositol 1-phosphatase synthase, or a mutant myo-inositol 1-				
CC	phosphate synthase having decreasing capacity for the synthesis for myo-				
CC	inositol-1-phosphate. Also included are a chimeric gene (comprising the				
CC	nucleic acid encoding soybean myo-inositol 1-phosphatase synthase, its				
CC	complement, subfragment or the complement of the subfragment, operably				
CC	linked to suitable regulatory sequences, where expression of the chimeric				
CC	gene results in a decrease in expression of an endogenous or native gene				
CC	encoding a soybean myo-inositol 1-phosphatase synthase), a soybean plant				
CC	comprising the chimeric gene (with a heritable phenotype of a seed phytic				
CC	acid content of less than 17 micromol/g, a seed content of raffinose plus				
CC	sucrose of less than 14.5 micromol/g, and a seed sucrose content of				
CC	greater than 200 micromol/g, provided that the plant is not LR33), seeds				
CC	from the plant, making a soybean plant with the heritable phenotype				
CC	(comprising crossing LR33 or the plant comprising the chimeric gene with				
CC	an elite soybean plant and selecting a progeny plant of the cross of				
CC	crossing step that has a heritable phenotype as mentioned above), seeds				
CC	of soybean plant made by the above method, a soy protein product derived				
CC	from seeds of a soybean plant (homozygous for one or more gene encoding a				
CC	mutant myo-inositol 1-phosphatase synthase having decreased capacity for				
CC	the synthesis of myo-inositol 1-phosphate, where the gene confers a				
CC	heritable phenotype as mentioned above), and making or producing a				
CC	soybean protein product derived from seeds of a soybean plant with a				
CC	heritable phenotype as mentioned above. The nucleic acid is useful for				
CC	altering raffinose saccharide, sucrose, phytic acid and inorganic				
CC	phosphate content of soybean seeds thus leading to valuable and useful				
CC	soybean products, since the presence of high concentration of raffinose				
CC	oligosaccharides in soy plants (and other legumes) can lead to flatulence				
CC	when consumed by humans. The present sequence encodes a mutant myo-				
CC	inositol 1-phosphatase synthase.				
SQ	Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;				
Alignment Scores:					
Pred. No.:	8,47e-251	Length:	1533		
Score:	2588.00	Matches:	501		
Percent Similarity:	99.02%	Conservative:	4		
Best Local Similarity:	98.24%	Mismatches:	5		
Query Match:	98.37%	Indels:	0		
DB:	13	Gaps:	0		
US-10-718-952-12 (1-510) x ADS81997 (1-1533)					
OY	1 MetPheIIeGIUAAnPhetysValGtusePerProAenVAllyrYTnGluThrgIuuile 20				
Db	1 ATGTTCATCGAGAAATTTAAGCTTGACGTCCTATATGGAAGTGCACCGAACGAGAATT 60				
OY	21 GlNSeVaTyTrAsntyrGrJuthrhGrGuuevUAIhsGiuaAnaGaEnGlThrTYr 40				
Db	61 CAATCGGTGACCAACAACGAAACCACGAAACTGTTCACGGAAACGGAATGGCACTCAT 120				
OY	41 GIATrrIlleValLyseProLysSerValAsntyGrGInPhelySTraSthrHiEvaIPro 60				

```
Db 121 CAGTGGATTGTCAACCCAAATCTGTCAATAGCAATTTAAACCAACATCCATGTTCTT 180
Qy 61 LysLeuGlyValMetLeuValGlyTyrGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AATATTAGGGGGTAAATGCTTGTGGGGTGGGAAACACGGCTCAACCCCTCAACGGTGGT 240
Qy 81 ValIleAlaAsnArgGluAspIleSerThrPalaThrLysAspLysIleGlnIleAlaAsn 100
Db 241 GTTATTGTGTCACCGAAGGGCAATTCATGGCTACAAAGACAAAGATTCAACAGCAAT 300
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTTTGGGTCCTCCCTCAACCAAGCTTACGCTATCCAGTTGGGTCTTCCAGGAGAGAA 360
Qy 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTATGCCCCCAATTCAGAGAGCTGCTTCCATGTGTAACTTGACGACATGTGTGGG 420
Qy 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GGATGGGATATCAGCAACATGAACTGCTGATGCTGATGCTGCAAGGCGCAAGGATGTTGAC 480
Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
Db 481 ATCGATTTCAGACAGACAGCTTGAGGCTTTCATGAAATTCATGCTTCACCTCCCGGAATC 540
Qy 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
Db 541 TATGACCCGGATTTCCTGCTGCTCAACCAAGAGAGCGGCCAACAGCTCATCAAGGCG 600
Qy 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysValPheLysGluAlaThr 220
Db 601 ACMAAGCAAGACCAAGTTCACAAATCATCAAAAGCATCAAGCGCTTTAAGAGAGCGCAC 660
Qy 221 LysValAspLysValValValLeuThrPalaAsnThrGluArgTyrSerAsnLeuVal 240
Db 661 AATGTCGACAGAGCTGTTTACTGTGACCTGCCACACAGAGAGGTACGTAATTTGGTT 720
Qy 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
Db 721 GTGGGCTTAAATGACCACTGAGGAATCTTGGGCTGCTGTGAGACGAATAGAGGCTGAG 780
Qy 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATTTCTCTTCACCTTGTATCCCATTTGCTTGTTATGAAATGTTCTTCTTCAATAT 840
Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATTTGATCTTGCCATTCGCGAGAACACT 900
Qy 301 LeuIleGlyGlyAspAspPheLysSerGlyInThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTGGTGGAGAGACCTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTGGTTGAT 960
Qy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleLeuGlyAsn 340
Db 961 TTCTCTTGGGGGGCTGTATCAAGCAACATCTATATGATCAATCAACCATCTGGGAAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db 1021 AATGATGGTATGAATCTTGGGCTCCACAAACTTTCCGTTCCAAAGAAATCTCCCAAGAGC 1080
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluIleAsp 380
Db 1081 AACGTTGTTGATGATATGCTCAACGCAATGCCATCTCTTAAGACCTCGTGAACATCCA 1140
Qy 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCAATGTTGTTGATTAATGATATGCTTACGTAAGGAGCAAGCAATAGAACCATGAT 1200
Qy 401 GUTYTTTTrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
```

```
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAAGAGACACCATTTGTTGCAACAACATGC 1260
Qy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGGATTCCTCTTGTAGCTGCTCTCTATTAATCTTGAGACTTGGTCTTGTGCTAGCTCAAC 1320
Qy 441 ThrArgIleGlnPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAGAAATCGAGTTTAAAGCTGAAATAAGGAGAAATTCACACTCATTCACCCAGTTGCT 1380
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATCTCAGCTACTCAGCAAGGCTCTCTGAGTTCCACGGGTACACCACTGATGAT 1440
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCATTGTCAAAGCAGCGTCAATGCTGGAATAAATATGAGGGCTTGTTGATGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1501 CCGAGAAATTAACATGATTTCTCGATPACAG 1530

RESULT 15
AAK90402
ID AAK90402 standard; cDNA to mRNA; 1950 BP.
XX
AC AAK90402;
XX
DT 24-SEP-1999 (first entry)
XX
DE Nicotiana paniculata INPS encoding cDNA.
XX
KW Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
XX water stress; resistance; ds.
XX
OS Nicotiana paniculata.
XX
PN JP1187879-A.
XX
PD 13-JUL-1999.
XX
PF 26-DEC-1997; 97JP-00359773.
XX
PR 26-DEC-1997; 97JP-00359773.
XX
PA (NIBS ) JAPAN TOBACCO INC.
XX
DR MPI; 1999-451546/38.
XX
P-PSDB; AAY24477.
XX
New INPS gene derived from Nicotiana genus plant - useful for conferring
PT resistance to water stress to plants.
XX
PS Claim 2; Page 6-8; 8pp; Japanese.
XX
CC The present sequence encodes Nicotiana paniculata inositol monophosphate
CC synthase (INPS), designated NpINPS1. INPS can be used to confer water
CC stress resistance to a plant
XX
SQ Sequence 1950 BP; 570 A; 377 C; 425 G; 578 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9,41e-239 Length: 1950
Score: 2470.00 Matches: 471
Percent Similarity: 96.47% Conservative: 21
Best Local Similarity: 92.35% Mismatches: 18
Query Match: 93.88% Indels: 0
DB: 2 Gaps: 0

US-10-718-952-12 (1-510) x AAK90402 (1-1950)
Qy 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyrThrGluThrGluIle 20
```


Db 92 ATGTTATTGAGAAATTTTAAAGGTGAGAGCCCAAGTTTAAAGTACCCGAAAGTGAATTT 151
QY 21 GlnSerVal1TyrAsnTyrGluThrGluLeuVal1HisGluAsnArgAsnGlyThrTyr 40
Db 152 CACTGTGCTAAGATTATCAAACTGAGATTAGTTCATGAGACAAATAAGGACATAT 211
QY 41 GlnTrp1LeuVal1ProLysSerVal1AsnTyrGlnPheLysThrAsnThrHisVal1Pro 60
Db 212 CAATGAGACCGTCAAGACCTTAAGCTGTCAAAATATGAGTTCAAGACTGATGTTCAATGTTCC 271
QY 61 LysLeuGlyVal1MetLeuVal1GlyTrpGlyVal1AsnAsnGlySerThrLeuThrGlyGly 80
Db 272 AAATTAGGGGTATATCTGTTGGATGGAGTGGAAAATAATGTTCAACCTTGACCGGTGTGT 331
QY 81 Val1Leu1AsnArgGluAsp1LeuSer1Trp1AlaThrLysAspLys1LeuGln1Asn 100
Db 332 GTTATGCTTAACAGAGAAAGAAATTTCAATGGCCCAAAAGATAGGTCAACAAACCAT 391
QY 101 TyrPheGlySerLeuThrGln1AlaSer1Ala1LeuVal1GlySerPheGlnGlyGlu 120
Db 392 TACTTTGGCTCTCTTACTCAGGCTTCTACTATTCAGATTTGGGTCTTTCATGGAGAAAG 451
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetVal1AsnProAspAsp1LeuAlaPheGly 140
Db 452 ATCTATGCTCCATTAAAGCCTCCTCCAAATGTCAAATCCAGATGAGTGTGTTTGA 511
QY 141 GlyTrpAsp1LeuSerAsnMetAsnLeu1AlaAsp1AlaMetAlaArgAlaLysVal1PheAsp 160
Db 512 GGATGGGACATGAGCAATGATTAAGCAATGACATGCAAGGAGGCTTAAGGTTATTTGAT 571
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetVal1ProLeuProGlyLe 180
Db 572 ATTGATCTTACAAAGACAGTTGAGGCCCTTACATGAGATCATGATGCCCTCCCTGGTATTC 631
QY 181 TyrAspProAspPhe1Leu1AlaAsnGlnGluGluArgAlaAsnAsnVal1LeuGly 200
Db 632 TATGACCTGATTTTCACTGCTGCTTAAACCAAGGTCACGTCACAAACGTCATCAAGGA 691
QY 201 ThrLysGlnGluGlnVal1GlnGln1LeuLysAsp1LeuVal1PheLysGlu1AlaThr 220
Db 692 ACCAAGAAAGAAACAATGATCAATCATTTAAGATATTAGGAGTTTAAAGAAAGAAC 751
QY 221 LysVal1AspLysVal1Val1LeuTrp1Thr1AsnThrGluArgTyrSerAsnLeuVal 240
Db 752 AAAGTGAACAAGGTGTGATATTGTGACTGCTTAACACTGAAGATACGTAATGTGTT 811
QY 241 Val1GlyLeuAsnAsp1ThrMetGluAsnLeuLeu1Ala1Val1AspArgAsnGlu1ArgL 260
Db 812 GTTGGACTTAATGACATATGAAAAACCTCTTGGCTTGTGTGACACAGAAATGAAGCTGAA 871
QY 261 IleSerProSerThrLeuTyrAla1Leu1CysVal1MetGluAsnVal1ProPhe1LeuAsn 280
Db 872 ATATCTCTCCACTTGTATGCTATGCTGCTATGCTTGAATAATGTGCTTTTATTAAT 931
QY 281 GlySerProGlnAsnThrPheVal1ProGlyLeu1Leu1AspLeu1Ala1Leu1ArgAsnThr 300
Db 932 GGAAGCCCCCAACAACCTTTGTCCAGGCTCATTTGATTTGCCATCAAGAAAGAACACA 991
QY 301 Leu1Leu1GlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerVal1LeuVal1Asp 320
Db 992 TTGATTTGGTGTGATGACTTTAAGAGTGTGTCAACCAAAATGAAGTCAGTGTGTGAT 1051
QY 321 PheLeuVal1GlyAlaGly1LeuLysProThrSer1Leu1SerTyrAsnHisLeuGlyAsn 340
Db 1052 TTCCCTGTGTGAGCTGTGATTAAGCCAAATCATATTTGAGCTACACACCATTTGGGTATAC 1111
QY 341 AsnAspGlyMetAsnLeuSer1AlaProGlnThrPheArgSerLysGlu1LeuSerLysSer 360
Db 1112 AATGATGGAATGATATCTGTCTGCCCTCAAACTTTCGGTCAAGAGAGATCTCGAAAAGT 1171
QY 361 AsnVal1Val1AspAspMetVal1AsnSerAsnAla1LeuTyrGluProGlyGlyHisPro 380
Db 1172 AATGTGTGTGATGACATGTGTTCAAGCAATGCCATCTTTATGAGCTGTGAGAGACACCT 1231

QY 381 AspHisVal1Val1LeuTyrVal1ProTyrVal1GlyAspSerLysArgAlaMetAsp 400
Db 1232 GACCATGTTGTGTGATTAAGTATGTGCATATGTGGAGACACAGAGGGCAATGGAT 1291
QY 401 GluTyrThrSerGlu1LeuPheMetGlyGlyLysAsnThr1Leu1LeuHisAsnThrCys 420
Db 1292 GAGTACACATCTGAGATTTTCATGTGGGGGAAAGAACACATTTGTTTGCACAAATACCTTGT 1351
QY 421 GluAspSerLeuLeu1Ala1Pro1LeuAspLeuVal1LeuLeu1AlaGluLeuSer 440
Db 1352 GAGATTCACCTTTTAACTGCTCCAAATTAATGATTTGGTCTTCTTGTGCAACTCAGT 1411
QY 441 ThrArg1LeuGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProVal1Ala 460
Db 1412 ACCCGCATTCAGCTCAAAAGCTGAAGAGAGAGGTAAATTCACCTCCTTCCACCCGTGGCT 1471
QY 461 Thr1LeuSerTyrLeuThrLysAla1ProLeuVal1ProProGlyThrProVal1Val1Asn 480
Db 1472 ACTATCTCAGCTACCTTAACCAAGGCTCTCTGTGTACCAACAGGTACACCAAGTGTGAT 1531
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsn1LeuMetArgAlaCysVal1GlyLeuAla 500
Db 1532 GCACCTCAAAAGCAAGAGGCAATGCTTGAGAACATATTGAGGGCTTGTGTGACTTCCA 1591
QY 501 ProGluAsnAsnMet1LeuGluTyrLys 510
Db 1592 CCAAGAAACAACATGATTCGGAATACAAA 1621

Search completed: June 7, 2005, 17:48:12
Job time : 671.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2005, 17:21:22 / Search time 3731.5 Seconds
(without alignments)
5202.406 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MFENFKYESNVKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2.1/USPTO_spool_p/US10718952/runac_06062005_173400_12856/app_query.fasta_1.1294
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdt -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEMAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952 @CGN 1 1.6628 @runac_06062005_173400_12856 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2356	89.5	1740	9	CL961092 OeIFCC005
c	2077	78.9	2360	1	AJ583520
2	1977.5	75.2	1828	3	CNS004MK BX827819 Arabidops
4	1606	61.0	1813	3	AK079323 AK079323 Mus muscu
5	1606	61.0	1814	3	AK005029 AK005029 Mus muscu
6	1561	59.3	988	7	CK279950 EST724028
7	1455	55.3	936	7	CK279064 EST725142
8	1438	54.7	972	7	CV135801 EST847010
9	1437	54.6	1613	3	CR599696 full-length

10	1427	54.2	914	7	CK275352	EST721430
11	1422	54.0	953	7	CO414034	EST844419
12	1391	52.9	898	7	CO113870	GR_Eb014
13	1375	52.3	865	7	CO085839	GR_Ea03E
14	1374	52.2	883	6	CD438045	BL01N0508
15	1374	52.2	891	7	CV290142	aofo1-10m
16	1367	52.0	906	7	CK320178	L2P08a02
17	1363	51.8	912	7	CV142333	EST853542
18	1359	51.7	867	6	CB973407	CAB30002
19	1357	51.6	860	7	CO072025	GR_Ea30P
20	1354	51.5	892	7	CO409215	EST839600
21	1347	51.2	813	7	CK768601	Gm-r1030-
22	1341	51.0	842	7	CF211061	CAB20007
23	1333	50.7	891	7	CF212628	CGF100434
24	1331	50.6	904	7	CV148446	EST859655
25	1335	50.4	836	7	CF518404	CAP007_1
26	1319	50.1	860	7	CO082095	GR_Ea45I
27	1304	49.6	960	7	CK274458	EST720536
28	1303	49.5	841	7	CO075472	GR_Ba36C
29	1285	48.8	830	7	CO107630	GR_Eb003
30	1280.5	48.7	1466	3	CR600268	full-length
31	1280.5	48.7	1513	3	CR606341	full-length
32	1273	48.4	831	7	CO132349	GR_EB45C
33	1262.5	48.0	834	7	CK933094	CGF100434
34	1260	47.9	970	7	CV147272	EST858481
35	1254	47.7	816	6	CB975215	CAB30006
36	1249	47.5	785	6	CB682609	OSJNFF10C
37	1248	47.4	878	7	CV137015	EST848224
38	1230	46.8	795	7	CO113224	GR_EB013
39	1228	46.7	800	7	CO112984	GR_EB004
40	1228	46.7	843	7	CK267514	EST713592
41	1222	46.4	739	4	BG124172	EST469818
42	1220	46.4	799	7	CF401745	RTW1_14
43	1209	46.0	783	2	BE130712	L48-91873
44	1208	45.9	779	6	CD448699	EKO7D2305
45	1207	45.9	851	7	CO129883	GR_EB31F

ALIGNMENTS

RESULT 1
LOCUS CL961092 1740 bp DNA linear GSS 21-SEP-2004
DERIVATION OeIFCC005463 Oryza sativa Expressed Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.

ACCESSION CL961092
VERSION CL961092.1 GI:52376905
KEYWORDS GSS.

SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 1740)
AUTHORS Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.

TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers
1..1740
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"

FEATURES

source

/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "

ORIGIN

Alignment Scores: 8,64e-276 Length: 1740
Pred. No.: 2356.00 Matches: 451
Score: 94.11% Conservative: 28
Percent Similarity: 88.61% Mismatches: 30
Best Local Similarity: 89.35% Indels: 0
Query Match: 9 Gaps: 0

US-10-718-952-12 (1-510) x CL961092 (1-1740)

Oy 1 MetPheIIeGluAsnPhelYsValGluSerProAsnValYsTyrThrGluThrGlu 20
Db 1 ATGTTTCATCGAGAGCTTCGCGTGGAGAGCCGCGACGTGGTACGGCGCGGAGATC 60
Oy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 61 GAGTGGAGCTACCACTAGACACGACGAGCTGGTGCCAGAGAGCCAGACGCGCCCTCC 120
Oy 41 GlnTrrIleValYsProLysSerValAsnTyrGlnPheLysThrAsnThrHisValPro 60
Db 121 CGCTGGTGTCTCCGCCCAAGTCCGTCGCTACACTTCCGACACACACCGTCCCC 180
Oy 61 LysLeuGlyValMetLeuValGlyTrrGlyValAsnAsnGlySerThrLeuThrGly 80
Db 181 AAGCTCGGGGTGATCTCGTGGGTGGGGCGGCAACCGCTCAACCTGACGCTCGG 240
Oy 81 ValIleAlaAsnArgGluAsnIleSerTrrPalaThrLysAspLysIleGlnGlnAlaAsn 100
Db 241 GTCATGCCCAACAGGAGGAAATCTCATGGCGCACCAAGACAGACAGTGCAGACCCAA 300
Oy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTATGGCTCATCAACCGAGCGTCCACCATCAAGGTCCGAGCTCAACCGGAGGAG 360
Oy 121 IleTrrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTAAGCGCCCTTCMAAGGCTCTCGCCCATGAGTGAACCCCTGATGACCTTGTTCGG 420
Oy 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GGCTGGACATTAGCAACATGAACCTGGCTGATGCTATGACCAAGGCGCAAGGTTACTT 480
Oy 161 IleAspLeuGlnLysGlnLeuArgProTrrMetGluSerMetValProLeuProGlyIle 180
Db 481 ATTGATCTGCAGAAACAGCTCAGACCTTACATGAGTCCATGATGCTCTCCCGGCATC 540
Oy 181 TyrAspProAspPheIleAlaAlaAsnGlnGluValArgAlaAsnAsnValIleLysGly 200
Db 541 TATGACCCCGACTTATCGCCGCCAACCAAGGATCCCGCGCAACATGCTCAAGAGGC 600
Oy 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db 601 ACCAAGAGAGACCAATGAGCAATCATCAAGACATCAAGGAATTCAGAGAAAGAGC 660
Oy 221 LysValAspLysValValValLeuTrrThrAlaAsnThrGluArgTrrSerAsnLeuVal 240
Db 661 AAAGGAGCAAGGTGTGTGTGTGGACTGCAACCACTGAAGATGACCAATCTCTGT 720
Oy 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaAlaValAspArgAsnGluAla 260
Db 721 GTTGGGCTCAATGACCAATGAGAACTCTCGGCTGTGTGCAACAAGAACAGCGGAG 780
Oy 261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATATACCATTCAGACCTGATGCAATGCTGTGTGTCATGAGAGGTATACCGTTACTT 840
Oy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300

Db 841 GCGAGTCTCTCAGAACACCTTTGTGCTGGGCTGATCGATCTTGCTATTAGAACATCTGC 900
Oy 301 LeuIleGlyIysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 CTGATGGTGTGTGATGATTTCAAGATGAGACAGACCAAGATGAATGCTGTCTTGTTGAT 960
Oy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 961 TTCCATGTTGTGCTGGAATTAAGCCCACTCAATTGTCAGTCAACACCTTGGGAGAT 1020
Oy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db 1021 AATGATGCGATGAACCTTTCGCACTTCMAACATTCGATCCATCAAGAGATATCCAAAGGC 1080
Oy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTrrGluProGlyGluHisPro 380
Db 1081 AATGTTGATGATGATGCTTCAAGCATCCATCTCTTATGAGCTCGGAGAGATCTT 1140
Oy 381 AspHisValValIleLysTrrValProTrrValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GATCATGTTGTGTGATCAATATGTGCGTATGTGGAGACAGAGAGGCAATGGAC 1200
Oy 401 GluTrrThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
Db 1201 GAGTACACCTCAGAGATCTTTCATGGGGGTGAAGACACATCGTTCGACAAACCTGT 1260
Oy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluSer 440
Db 1261 GAGGACTACTTCTTCCGCCGCCATCATCTTATCTTGCTGCTCTTCCGAGCTCAGC 1320
Oy 441 ThrArgIleGlnPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACCAGGATTCAGCGGAAGCCGAGGGGAGAGAGAGTTCATTCTTCCATCCAGTGCT 1380
Oy 461 ThrIleLeuSerTrrLeuThrLysAlaProLeuValProProGlyTrrProValAlaAsn 480
Db 1381 ACCATCTCGAAGCTCTCCACCAAGGCACTCTTGTCTCTCGCACACCATGTGTGAC 1440
Oy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCCCTGGCAAGACAGAGGCAATGCTTGAGAACATCATGAGGCTCGTGGCTGACC 1500
Oy 501 ProGluAsnAsnMetIleLeuGluTrr 509
Db 1501 CCCAGAACCAATGATCTTGAGTAC 1527

RESULT 2
AJ583520/c 2360 bp mRNA linear EST 08-JAN-2004
LOCUS AJ583520 PpProt1 Physcomitrella patens cDNA clone pp001015066
DEFINITION similar to putative myoinositol-1-phosphate synthase, mRNA
sequence.
ACCESSION AJ583520
VERSION AJ583520.1 GI:40781590
KEYWORDS EST.
SOURCE Physcomitrella patens
ORGANISM Physcomitrella patens
REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
1 (bases 1 to 2360)
Hohn, A., Egner, T., Lucht, J.M., Holtorf, H., Reinhard, C., Schween, G.
TITLE An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, Physcomitrella patens
JOURNAL Curr. Genet. 44 (6), 339-347 (2004)
COMMENT Plant Biotechnology
Contact: Schween G
University of Freiburg
Sommerstrasse 5, D-70104 Freiburg, Germany.
FEATURES
source 1..2360
/organism="Physcomitrella patens"

```

/mol_type="mRNA"
/db_xref="taxon:3218"
/clone="pp001015066"
/issue_type="Protonema"
/clone_1ib="ppProci"

```

ALIGNMENT SCORES:

```

Pred. No.: 1.6e-241 Length: 2360
Score: 2077.00 Matches: 400
Percent Similarity: 88.61% Conservative: 59
Best Local Similarity: 77.22% Mismatches: 51
Query Match: 78.94% Indels: 8
DB: 1 Gaps: 3

```

US-10-718-952-12 (1-510) x AJ583520 (1-2360)

```

QY 1 MetPheilegIuaAenPhelysValGluSerProAenValYerThrGluThrGluile 20
DB 2150 ATGTTGTCGACAGCTTTCGTTGAGAGTCTCTCATGTGATACACGAGACATGATC 2091
QY 21 GlnSerValTYrAsnTYrGluThrThrGluLeuValHsGluAenArg----AsnGly 38
DB 2090 AAGTCAGCTATGAGTACCAACACCGAGTCTCTGACAGACGACGACGACGATGCG 2031
QY 39 ThrTYrGlnTPileValYerProLYserValAsnTYrGlnPhelyThrAsnThrHis 58
DB 2030 AAGTATGAGTGGTGGCCAGGACCAAGTCTGTCATGCAATTCGCACTGACGCGCAAG 1971
QY 59 ValProlYsLeuGlyValMetLeuValGlyTPGlyGlyAsnAsnGlySerThrLeuThr 78
DB 1970 GTGCTTAAGCTGGAGATGATGCTCGTGGGTGGGTCACACGAGTCCACTCTCAG 1911
QY 79 GlyGlyValIleAlaAsnArgGluAspIleSerTPAlaThrLYsAsp-LysIleGlnG 98
DB 1910 GCTGATATCTCGCCAAAGAGATATCTCATGAGTGCACCAAGATGCGGTGACAGA 1851
QY 98 nAlaAsnTYrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySer-PheGln- 117
DB 1850 CGCCAAATTAATCTGAGATCTCTACACAGCTTGGCGCGGTGGGTCTCTTTCACAG 1791
QY 118 --GlyGluGluIleTYrAlaProPhelySerLeuLeu---ProMetValAsn-ProAsp 135
DB 1790 GTGAAAGAAAGATTAATGATCTTTAAAGACATTTACACGATGTCACACCCCAAC 1731
QY 136 AspIleValPheGlyTYrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArg 155
DB 1730 GATCTGATGTCGCGGTGGGATATCTCCAACTGAACTTGGCGGATGCGATGATGCT 1671
QY 156 AlAlaYsValPheAspIleAspLeuGlnLYsGlnLeuArgProTYrMetGluSerMetVal 175
DB 1670 GCGCGCGTGTGAGATGATCTTACAGACCCCACTTCATCGCTGACACAGAGGCTGTGCAAC 1551
QY 176 ProLeuProGlyIleTYrAspProAspPheIleAlaAsnGlnGluLysArgAlaAsn 195
DB 1610 CCCATGCTGTGATTTACAGACCCCACTTCATCGCTGACACAGAGGCTGTGCAAC 1551
QY 196 AsnValIleLYsGlyTYrAspGlnGluGlnValGlnGlnIleIleLYsAspIleLYsAla 215
DB 1550 AATCTCATCAAGGCTTCCAGAGAGCAAAATGAGCGGATTTATCAGACATACGGAAC 1491
QY 216 PheLYsGluAlaThrLYsValAspLYsValValIleUTPThrAlaAsnThrGluArg 235
DB 1490 TTCAAGGCGCAAAACCAAGGTGACAAAGTAGTACTGTGACCGCTTACACAGAGGCT 1431
QY 236 TYrSerAsnLeuValValGlyLeuAsnAspThrMet-GluAsnLeuLeuAlaValAs 255
DB 1430 TACAGCGAGTATGTGAGTGTGATGACACCAAGGAATCTGTTGATCTCTGGA 1371
QY 255 pArgAsnGluIaGluIleSerProSerThrLeuTYrAlaIleAlaCYsValMetGluAs 275
DB 1370 GAAGGAGAAATTAATCTGCGCGTGCACCTTGTATAGCTCTGCTTGCATCAGAGACA 1311

```

```

QY 275 nValProPheileAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAl 295
DB 1310 GATTCCTTCATCAACGCGAGTCCCAAGACATTCCTGCTGAGTATTAATGAAATGCC 1251
QY 295 alIleAlaArgAsnThrIleuIleGlyLYsAspAspPheLYsSerGlyGlnThrLYsMetLY 315
DB 1250 TGTGAAAGAACTCATTAATTTGTTGTGACGACTTCAAGACGCGTCAAAACAGATGAA 1191
QY 315 sSerValLeuValAspPheLeuValGlyAlaGlyIleLYsProThrSerIleValSerLY 335
DB 1190 ATCCCTTCTAGGACTTCTTATGTTGAGCGGGAATTAAGCCAAAGCTCAATTTAGTCA 1131
QY 335 rAsnIleuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLY 355
DB 1130 CAACCATTTGGGCAACCAAGATGATGATCTTCCGCCCCCTCAACATTCCTTCCAA 1071
QY 355 sGluIleSerLYsSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTYrG 375
DB 1070 GGAATCTTCCAAAGACCAATGTATGTGACGACATGTTGGCTGACAAATTCATTAACGC 1011
QY 375 uProGlyGluIleProAspPheValValIleLYsTYrValProTYrValGlyAspSe 395
DB 1010 GCCCGGTGAACATCCCGACCAAGTATTGTGATCAAGTACGCTCTTATGTGGGTGACAG 951
QY 395 rLYsArgAlaMetAspGluTYrThrSerGluIlePheMetGlyLYsAsnThrIleVal 415
DB 950 CAAGAGGCGAATGACGATGATCACTTGCAGATCTTCATGAGCGGACGATCAACATCGT 891
QY 415 lLeuHisAsnThrCYsGluAspSerLeuLeuAlaIleAlaIleLeuAspLeuValLe 435
DB 890 CATGACATATAGTGCAGAAATCTTTACTGCGCGCGCGCTCACTTGGACTTGTGCT 831
QY 435 uLeuAlaGluLeuSerThrArgIleGlnPheLYsAlaGluAsnGluGlyLYsPheHisSe 455
DB 830 GCTGCTAGCTCTGCACAGATCAATTCAGCAAGAACAGACGAGAAATTCATTC 771
QY 455 rPheHisProValAlaThrIleLeuSerTYrLeuThrArgLYsAlaProLeuValProPro 475
DB 770 TTTCACCGCGGTGCTGATCTCTGATCACTTCAACAGGCTCCCTGTGATCCCCACAG 711
QY 475 TYrProValValAsnAlaLeuSerLYsGlnArgAlaMetLeuGluGluAsnIleMetArgAl 495
DB 710 TACCCCAATGTGTTACGCGTTCGTCACAGAGGCGCGATGCTAGAGATATCATGCGCG 651
QY 495 aCYsValGlyLeuAlaProGluAsnAsnMetIleLeuGluTYrLYs 510
DB 650 GTGATGTGGGTGTCGCCGCAATTAACATGATGCTAAGATCAAG 605

```

RESULT 3
CNS0A4MK 1828 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSURP3242E06 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
BX827819.1 GI:42462602
ACCESSION BX827819.1
VERSION HTC; GSUR_cDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1828)
Castelli,V., Aury,J.M., Jallion,O., Winkler,P., Clepet,C.,
Memard,M., Gruand,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
JOURNAL 2 (bases 1 to 1828)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE

JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads. Life technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castel11 V., Aury J.M., Jailion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length

http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.

location/Qualifiers

1. 1828

organism="Arabidopsis thaliana"

molecule="cDNA"

strain="Col-0"

db_xref="taxon:3702"

clone="GSLRPH24Z06"

tissue_type="Hormone Treated Callus"

plasmid="pCMVSPORT 6"

complement(1. 1828)

gene "At4g39800"

ORIGIN

Alignment Scores:

Pred. No.: 1,51e-229 Length: 1828

Score: 1977.50 Matches: 397

Percent Similarity: 82.55% Conservative: 43

Best Local Similarity: 74.48% Mismatches: 52

Query Match: 75.16% Indels: 41

Gaps: 6

DB: 3

US-10-718-952-12 (1-510) x CNS0A4MK (1-1828)

QY 1 MetPheilegIuAnPheleValGluSerProAnValIyTyThrGluThrGluile 20

Db 58 ATTTTATTGAGAGCTTCAAGTTGAGAGCCCGAAGTGAAGTACAGAGAGTATGATTT 117

QY 21 GlnSerValIyAsnTyGluThrThrGluLeuValHISgluAsnArg--AsnGlyThr 39

Db 118 CACTCTGTCAGATTACGAGACGAGGTTGTTCAATGAGAAACCCGTCAATGSCACT 177

QY 40 TyrgIntrPileValIyProIySerValAsnTyrgInPheIyThrAsnThrhISVal 59

Db 178 TACCAATGAGATTGTGACCAAGACCTGTCAAAATGCAATTCAAACCGATTCGTGTC 237

QY 60 ProIySleuGlyValMetLeuValGlyTrpGlyGlyAsnAsnGlySerThrIleuThrly 79

Db 238 CCCCAATTAGGGGTTATCTGTGGGTTGGAGGAAACAAATGATCAACTCTCACCGCT 297

QY 80 GlyValIleAlaAsnArgIuAspIleSerTrpAlaThrIyAspIyIleGlnGlnAla 99

Db 298 GGTGTGATTGCGAACAAGAGATCTGTGGGACGAGAGCAAGAGCAAGTGCACAGCG 357

QY 100 AsnTyPheGlySerIleuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyIu 119

Db 358 AATTACTTCGGGTCACTAATCACTCATGATTCGTGTCGATCTTTTAAAGGTGA 417

QY 120 GluIleTyAlaProPheIySerIleuLeuProMetValAsnProAspAspIleValPhe 139

Db 418 GAGATCTATGCTCTTCAAGAGTCTCTTCCAAATGGAATCCGATGATCTTGTGTT 477

QY 140 GlyIyTPAspIleSerAsnMetAsnIleAlaAspAlaMetAlaArgAlaIyValPhe 159

Db 478 GAGGATGGGATATAGTACATGAACTTACGAGATGCCATGGCTTACAGCCGGGTTCTT 537

QY 160 AspIleAspLeuGlnIyGlnIleuArgProTyrmGluSerMetValProIeuProGly 179

Db 538 GATATCGACTTGAGAAACAGCTCAGGCTTACATGAGAAACATTTGCCATCTCCCTGG 597

QY 180 IleTyAspProAspPheIleAlaAsnGlnGluIuArgAlaAsnValIleIyS 199

Db 598 ATCTTCATCTGATTTCATAGCTGCTTATCAAGGTCACGCTTACCAAGTATCAAA 657

QY 200 GlyThyrgInGlnIuValGlnGlnIleIleIySAspIleIySAlaPheIyGluAla 219

Db 658 GGTACCAAGAGAAACAAGTGCACCATCATCAAGACATGAGGAGGATTAAAGAGAG 717

QY 220 ThrIySValAspIyValValIleuThrPheAlaAsnThrGluAlaIyTySerIleu 239

Db 718 AATAGAGGTAAAGTGTGCTCTGAGCGGCTTACACAGCGCTTATAGCAATGTG 777

QY 240 ValValGlyIleuAsnAspThrMetGlnAsnIleuAlaAlaValaAspArgAsnGluAla 259

Db 778 GTCTGGGATGAAACGATACATGAGATTCATGAGTCTGTGATAGGAGATGAGGCT 837

QY 260 GluIleSerProSerThrIleuTyAlaIleAlaCySValMetGluAsnValProPheIle 279

Db 838 GAGATCTCTCTTCAACCTTATGCGATTCCTGTGTCTTGAAGGAAATCCGTTCTAC 897

QY 280 AsnGlySerProGlnAsnThrPheValProGlyIleuIleAspIleAlaIleAlaArgAsn 299

Db 898 AATGAGAGCCCTCAAAACCTTTGTTCCAGTCTAATGATTAAGGCGATGAGAAACAT 957

QY 300 ThrIleuIleGlyIyAspAspPheIySerGlyInThrIySMeIySerValIleuVal 319

Db 958 GTTTGATCGGTGAGATGATTTCAAGAGTGTCAAAACCAAGATGAAATCTGTGTT 1017

QY 320 AspPheIleuValGlyAlaGlyIleIySProThrSerIleValSerTyAsnIleuGly 339

Db 1018 GATTTCCTAGTTGGCGGAGATCAAGCTCACTCAATTGAGACTCAATCACTTGGGA 1077

QY 340 AsnAsnAspGlyMetAsnIleuSerAlaProGlnThrPheArgSerIyGlnIleSerIyS 359

Db 1078 AACACAGATGAAATGAAATCTCTCAGCTCCACAGACTTTCAGATCTTAAGGATCTCCAA 1137

QY 360 SerAsnValAlaAspAspMetValAsnSerAsnAlaIleIyTyrgIuProGlyGluHIS 379

Db 1138 AGCATATGCTGAGACGATATGTTGCTAGCATGATCTCTTTCAGACTCGGGGAAC 1197

QY 380 -ProAsp-HISValValIleIle-----IyTyValProTyValGlyAspSerIyS 397

Db 1198 CTGGACCCATGTTGTTCTATCAACAGAGAAACATGACACATCATCAAGACATGA 1257

QY 397 TG-----AlaMetAspG 401

Db 1258 GCGAGTTTAAAGAGAAATAGGTGATTAAGTTGTTCTCGACGCGCTTAACACAG 1317

QY 401 IuTyTrhSerGluIlePheMetGlyGlySAsnThrIle-----ValIeuHISAsnT 419

Db 1318 AGCGTTATAGCAATGTGCTGTGGATGAACATCAATGAGAAATTCATGAGAGTCTG 1377

QY 419 hrCySgluAspSerIleuIleAlaProIleIleIySAspIleValIleuIleuIleGlu 439

Db 1378 TGATAGGATGAG-----GCTGAGA 1398

QY 439 euserThrArgIleGlnPheIySalagluAsnGluGlyIySAsnThrhISerPheHISProIy 459

Db 1399 TCTCTCTTCAACCTTAT-----GGAAAGTTTCACTTCTTCACTCTCTG 1443

QY 459 AlaIleThrIleIeuSerTyTrhIleuThrIySAlaProIeuValProProGlyThrProValI 479

Db 1444 TGGCCACCATCTACGACTCAACCAAGGACCCCTTGTCGCGCGGAACTCCGGTGA 1503

QY 479 AlaAlaIleuSerIySAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 499

Db 1504 TTAACGAATCTGCAAGAGCGGCTATGCTGAGAAATATATATAGAGGCGTGCCTTGAC 1563

QY 499 euaIaProGluAsnAsnMetIleuGluTyTyIyS 510

Db 1564 TGGCTCCGAGAAACAATGATCATGAAATCAAG 1598

RESULT 4	AK079323	1813 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK079323				
DEFINITION	Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630023013 product:MYO-INOSITOL-1-PHOSPHATE SYNTHASE A1 (EC 5.5.1.4) homolog [Homo sapiens], full insert sequence.				
ACCESSION	AK079323				
VERSION	AK079323.1	GI:26347916			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionahni; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Kitanaka, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsumi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
REFERENCE	5				
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.				
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
JOURNAL	Nature 420, 563-573 (2002)				
REFERENCE	6 (bases 1 to 1813)				
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashizaki, K., Hayatsu, N., Hizumoto, K., Hirooka, T., Hirose, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Ohashi, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takai-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Yamatsuta, M., and Hayashizaki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222,				

COMMENT

Fax:91-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/
 Location/Qualifiers

FEATURES

source	1 . . . 1813 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:9630023013" /db_xref="taxon:10090" /clone="9630023013" /issue_type="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="16 days neonate" 50 . . . 1723 /note="unnamed protein product; MTO-INOSITOL 1-PHOSPHATE SYNTHASE A1 (BC 5.5.1.4) homolog [Homo sapiens] (SPTR1Q9NPB2, evidence: FASTA, 88.4%ID, 100%length, match=1668) putative" /codon_start=1 /protein_id="BAC037607.1" /db_xref="GI:26347917" /translation="MEPAAEILVDSPDVYSGPETEARERYRTTRVSRREGVLAVQO ARFPFRATROVPRLGVNIVGKGNNGSTLVAAVIANRLTLWPRGRKANYYGGI TOAGYNLGLDENGREVRYPFBALLPKYAPNDLPFGMDISLIQAEMRAQVLD LDEQLPMHESLRPBPVSYLEPIANAOTAADNLIPETRAQQLERKDIIDRRSS GDLYKVTMTANTARPFCEVPERNDTANLTLLIQLEGSPTLFVAASIIEDCA NSGPDTVPGALIELASORHVFVGDDPSQTKSVYLVDPLIGSGIKTMSIVSYI LNQNDOINLSAPLOFRSKVEVTSVVDDMVSHNYLVARGBRPCHVIRKVPVYG KQALEDRSELMAGGTNTLYLNTCEDSLAAPIMDLVYLTELCOHYRPTSDDP QGHGYVLISLFLFKAPLVPPOSPVYNALIFQRSCIENTIFRACTGCLPPONMILEHI ERGEPRIKGEVVAVSPILCKKEPTPATNGCTGDANGHPQAPTEPLSTA" /notes=".1800 /size=".1800 /type="putative" 1813 /note="putative" 1813
--------	--

CDS

polyA_signal

polyA_site

ORIGIN

Alignment Scores:

Pred. No.:	3,36e-184	Length:	1813
Score:	1606.00	Matches:	303
Percent Similarity:	78.17%	Conservative:	91
Best Local Similarity:	60.12%	Mismatches:	104
Query Match:	61.04%	Indels:	6
DB:	3	Gaps:	4

```

US-10-718-952-12 (1-510) x AK079323 (1-1813)
Qy      8 ValGluSerProAsnValLyStYrThrglunhrgluiEginSeValryAsnTyrglu 27
       ++++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|+++++|
Db      74 GTGGATTAGCCCGGAGCTGTCTACAGCCCCGAACAATCGAGCGCGCTACGAGTCGG 133
Qy      28 ThTrtGluleuValHisgluAsnArgasnGlyThrTyrgIntPrIlleVallyeProlys 47
       |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      134 ACAACCGCGGTCGCCGCAG-----GGCGCGCTGCTGCGG-----GTCCAGCCCAAG 181
Qy      48 SerValAsnTyrcInphelyThrAnthrHisValProlySleugliValMetLeuVal 67
       :::   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      182 GCTACCGCGTTTACCCTTCGCCGACCGCGCGGACGGTGCCCACTCGGGGTATTGGTC 244
Qy      68 GlyTYrGlYgLYAsnAsnGlySerThrlEurThrglygiValilEalaasnarGluasp 87
       :+++:|+|++|++|++|++|++|++|++|++|++|++|++|++|++|++|++|++|
Db      242 GGCTGGGGGGGGAACAACGCTCCAACGCTCACTGCTGCTTGGCCCAATCGGCTGGCC 300
Qy      88 lIsErTrPalalThrlyehaplsylleEgnlnlaiaenTyrrheGlySerLeuThGln 107
       :+++++|::|++|++|++|++|++|++|++|++|++|++|++|++|++|++|++|
  
```

D	302	CTAACCTGCGCCACCGCAGCAGTGCAGGAAGGAGCAACTATATGATCGTTGACCCAG	361
O	108	AlAseraIaIleayValgIy---SerPheingIygluIuIeTyraIaProPhelys	126
D	362	GCGGGCACCTGGATCTGGATAGAAACGCGCGGAGAGTGTTCCTTCACT	421
O	127	SerLeuauProMetValAsnProAspIleValPheglYglYTrpAspIleSerAsn	146
D	422	GCGCTGCTAACCCATGATGTGCCCCCAACACCTGGTGTGATGTGGAGATATCTCCGTCG	481
O	147	MetAenLeuAlaAspAlaMetAlaArgAlaIyValPheAspIleAspLeuGlnIySgl	166
D	482	CTGAACCTGGCCGAGCCATGCGGCGCGCAGAGCTCTGAGCTCGGCTCTGAGGAAACAG	541
O	167	LeuArgProTyMetGluSerMetValProleuProglYIleTyraSpProAspPheIle	186
D	542	CTGTGCGCCCACTGAGAGCCTGGGTCCGCGGCGCTCAGCTCAATCTGAGTTCAATC	601
O	187	AlaAlaAsnGlnGluIuArgAlaAsnMetValIleIySglYThrIySglGlnGlnIuAl	206
D	602	GCTGCAACCAAGACAGACGTCGGGACAACTCATCTCGGACACGCGCCAAACAGTTG	661
O	207	GlnGlnIleIleIyAspIleIySglAlaPheIySglIuAlaThrIyValAspIlyValVal	226
D	662	GAGCAAAATCCGAAGGACATTAGAATTTCCATCCAGTCGGGAGTTGGATGAATCATC	721
O	227	ValIeuTrpThrAlaAsnThrGluArgTySerAenLeuValIaGlyLeuAsnAspThr	246
D	722	GTGCTGTGAGACGGCAATACGAGCGCTTCTGCGAGGTGTCCTCCAGTCCGAAATGACA	781
O	247	MetGluAenLeuLeuAlaIaValaAspArgangIuAlaGluIleSerProSerThrIeu	266
D	782	GCAGAAACCTGCTACATCAATACAGCTTGGC---CTGAGAGTCCACCCGTCCACACTT	838
O	267	TyralaIleAlaCyValMetGluAenValProPheIleAsnGlySerProGlnAsnThr	286
D	839	TTTGTCTGGCCAGCATCTCGAGAGACTGCGCTTCTCATGATCTCCACAGAACACA	898
O	287	PheValProglYLeuIleAspPheAlaIleAlaArgAsnThrLeuIleGlyIyAspAsp	306
D	899	CTGGTACCGGTCCTGGAACTGCTTCGACGCGCANTGTGTTGATGGTGTGATGAC	958
O	307	PheIySglYgluIuThrIyMetIySglSerValIleuValAspPheLeuValIaGly	326
D	959	TTCAAGTCAGGCAACATCAAGGTCAAGCTGTCCTGGTGACCTTCTCATCGGCTCTGG	1018
O	327	IleIySgProThrSerIleValSerTyraAsnIleSleuGlyAsnAsnAspGlyMetAenIeu	346
D	1019	CTCAAGACCAATCTCATGTGATGCTATTAACACTTGGGACAAACAGCGGCAACACTG	1078
O	347	SerAlaProGlnThrPheArgSerIySglIleSerIySerAsnValValaAspAspMet	366
D	1079	TCTGCACCGCTGCAGTTCGCGTCCAAAGAGGAGCAAAAGAGAGTGTGGAGCAATG	1138
O	367	ValaenSerAsnAlaIleLeuTyrgIuProglYgluIuIleSproAspHisValaValIle	386
D	1139	GTTCAACAAACCATGTCTCTACCGCGCTGGAGGCGGCAACCATGTGTGTGTATC	1198
O	387	LysTyraIaProTyValaGlyAspSerIySglAlaMetAspGluTyThrSerGluIuIle	406
D	1199	AAATATGTGCCCTATGTGGCGACACAGAGCGCTTGAGAGATCAACCTCCGAGCTG	1258
O	407	PheMetGlyIyIySglAsnThrIleValIleuHisAenThrCySgluAspSerLeuLeuAla	426
D	1259	ATGCTGGGTGGGACAAACACTTGTGTCTCATATACTGAGAGATTCGTCTCTGGCC	1318
O	427	AlaProIleIleLeuAspIeuValIleuAlaGluIeuSerThrArgIleGlnPheIyS	446
D	1319	GCGCCCATCATGCTGACCTAGTCTGCTCAACAGAGCTGTGTCAGCGGAGGCTTTCG	1378
O	447	AlaGluAsnGluGlyIyPheHisSerPheHisProValAlaIaThrIleLeuSerTyIeu	466
D	1379	ACAGACTCGAACCCGACCTCAGGCTTCCACAAAGTGTGTCTCTCTTAAAGTCTCTG	1438

OY	467	ThrlYsa1aP-roLcUeVn1P-roP-roGlyThP-roVa1Va1aSn1a1eUsErLySg1nArg	486
Db	1439	TTTAAAGCCCCCGCTTGTGCCCCCGGAGCGCCGTGATGTGATGCCCTTCCCGCAGCGC	1498
OY	487	AlaMetLeuGlUaSn1eMeArGa1aCySa1e1G1YleuN1a1P-roGlu1aSn1eMe11e	506
		::: ::::	::: ::::
Db	1499	AGCGTATCGAGATATTTTCAGGGCGTGGCGTGGGGCTCCCGCCCGACAGACATGCTA	1558
OY	507	LeuGlUTrYrLyS	510
Db	1559	TTAGAGCAACAAG	1570
RESULT 5			
AK005029			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			

QY 367 ValaSenSerAsnAlaIleuTyTgUpProGlyGluHisProAspHisValValIle 386
Db 1140 GTTACAGCAACCACTGCTTACCGCTCGAGAGCGGCCAGACACTGTGTGTGATC 1199
QY 387 LysTYValAlProTyTValGlyAspSerLysArgAlaMetAspGluTyThrSerGluIle 406
Db 1200 AATATGTGCTTATGTTGGCAGACAGCGCTGCTTGAGCAAGTACCTCCAGCTG 1259
QY 407 PheMetGlyGlyLysAsnThrIleValIleuHisAsnThrCysGluAspSerLeuVal 426
Db 1260 ATGCTGGGGGAGCAACACTTGGTGTCCATATATACCTCGAGAGATTGCTCGTGGCC 1319
QY 427 AlaPProIleIleLeuAspLeuValIleuValaGluLeuSerThrArgIleGlnIleHis 446
Db 1320 GGGCCCATCATCTGAGCACTGATGTGCTGCACAGAGCTGTGTCAAGCGCTGCTTGC 1379
QY 447 AlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyIleu 466
Db 1380 ACAAGACTCGAGCCCGAGCGCTCAGGGCTTCCACACAGTGTCTCCCTGTGACTTCTG 1439
QY 467 ThrIlyAlaProLeuValProProGlyThrProValIValAsnAlaLeuSerLysGlnArg 486
Db 1440 TTAAAGCCCGCGCTGTGCGCCCGGCGAGCCCTGTATGTAATGCTCTTCCGCAAGGCC 1499
QY 487 AlaMetLeuGluAsnIleMetArgAlaCysValaGlyLeuAlaProGluAsnMetIle 506
Db 1500 AGCTGTATCGAGAAATATTTTCAGGGCTTGCGTGGGGCTCCCGCCAGAACCAATGTCA 1559
QY 507 LeuGluTyTyrLys 510
Db 1560 TTAGAGCACAG 1571

RESULT 6
CK277950 988 bp mRNA linear EST 03-AUG-2004
LOCUS EST2724028 potato abiotic stress cDNA library Solanum tuberosum cDNA
DEFINITION clone POAB302 5' end, mRNA sequence.
ACCESSION CK277950
VERSION CK277950.1 GI:39834928
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum

REFERENCE
AUTHORS Bukeriyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
JOURNAL asterids; lamiales; Solanales; Solanaceae; Solanum.
COMMENT 1 (bases 1 to 988)
Buell C.R., Hart A., Zismann V., Karameyheva S.A. and Baker B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: EST724029
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potatoc-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/ .
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1. 988
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB302"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/note="vector: PCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at

following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots:1d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	4,08e-179	Length:	988
Score:	1561.00	Matches:	301
Percent Similarity:	95.44%	Conservative:	13
Best Local Similarity:	91.49%	Mismatches:	15
Query Match:	59.33%	Indels:	1
DB:	7	Gaps:	0

US-10-718-952-12 (1-510) x CK277950 (1-988)

QY 55 ThrAsnThrHisValProLysLeuGlyValMetLeuValGlyTyTgIlyValAsnAsnGly 74
Db 2 ACTAGTGTCAATGTCCAAA-TTAAAGGATTATGCTGTGTGATGGGGGAGAAACAATGTG 60

QY 75 SerThrLeuThrGlyGlyValIleAlaAsnArgGluAspIleSerThrAlaThrLysAsp 94
Db 61 TCAGCTTGACTGAGAGGTATTATGCTTACAGAAAGAAATTTCATGGGCTACCAAGAT 120

QY 95 LysIleGlnGlnAlaAsnTyPheGlySerLeuThrGlnAlaSerAlaIleArgValGly 114
Db 121 AAAGTCAGCAAGCAATTACTTGGCTCTTACACAGGCTCTACTATTGCGATTGGT 180

QY 115 SerPheGlnGlyGluGluIleTyTAlaProPheLysSerLeuLeuProMetValAspPro 134
Db 181 TCCTTCATAGAGAGAGATATATGCCCCCTTAAAGCAATATCCCATGGTCAATCCA 240

QY 135 AspAspIleValPheGlyTyTAspIleSerAsnMetAsnLeuAlaAspAlaMetAla 154
Db 241 GATGATGTAGTGTGGAGATGGGACATCGACATGAATTTGGCAGATGCCATGGCA 300

QY 155 ArgAlaIysValPheAspIleAspLeuGlnLysGlnLeuArgProTyTMetGluSerMet 174
Db 301 AGGGCTAAAGTTTGTGACATGATCTTCAAAAGCAGCTGAAGCCCTCATGGAATCCATG 360

QY 175 ValProLeuProGlyIleTyTAspProAspPheIleAlaAlaAsnGlnGluArgAla 194
Db 361 GTCCACACGCTGGTATCTATGACCCTGATTCATGTGTGCAACCAAGGCTCAGTGGCT 420

QY 195 AsnAsnValIleLysGlyThrLysGlnGluGlnAlaGlnIleIleLysAspIleLys 214
Db 421 AACATATGATTAAGAGAACCAAGAAACAAATTCATCAAACTTAAGATATTAGG 480

QY 215 AlaPheLysGluAlaThrLysValAspLysValValIleuTyTProAlaAsnThrGlu 234
Db 481 GAGTTTAAGAGAACCAACAAAGGTGATGTTGTTGTGAGCTGATCACTGAA 540

QY 235 ArgTyTSerAsnLeuValaGlyLeuAsnAspThrMetGluAsnLeuAlaAlaVal 254
Db 541 AGATACAGCACTGTGGCTGTGGCTTAATCATCTCAATGAAAACTTTGTCTTGTG 600

QY 255 AspArgAsnGluAlaGluIleSerProSerThrLeuTyTAlaIleAlaCysValMetGlu 274
Db 601 GATGAATAGGCTGAATATCTCTTCCACTTGTATGCTATTGCTGATTCTTGAA 660

QY 275 AenValProPhe1leAenGlySerProGlnAenThrPheValProGlyLeu1leAenLeu 294
 |||||
 Db 661 AATGTCCTTCATCATGAGAGCCCAAAACACTTTTGTCCAGGCTCATGATTG 720
 |||||
 QY 295 A1a1leA1aAAsnThrLeu1leGlyGlyAAspAAspPheLeuSerGlyGlnThrLyMet 314
 |||||
 Db 721 GCATCAACAGAGAACACTTGTGTTGTTGATGATCACTTAAAGTGTCAAAACAAGATG 780
 |||||
 QY 315 LysSerValLeuValAAspPheLeuValGlyA1aGly11leLysProThrSer1leValSer 334
 |||||
 Db 781 AAGTCAGCTCTGTGATTTCTCTGTGAGCTGTGATTAACCAACATCAATTTGAGAC 840
 |||||
 QY 335 TTAAsn1leLeuG1YAAsnAAspGlyMetAAsnLeuSerA1aProGlnThrPheArgSer 354
 |||||
 Db 841 TCAACACATTTGGGTACATGATGAAATGATCTTCGGCCCTCAGACTTTCCGCTCA 900
 |||||
 QY 355 LysGln1leSerLysSerAAsnValValAAspAAspMetValAAsnSerAAsn1a1leLeu1Yr 374
 |||||
 Db 901 AAGGAGATCTCAAAAGCAATGTTGTCGATGACATGCTTTCTAGTAATGCCATCTTTAT 960
 |||||
 QY 375 GluProGlyVal1leH1aProAAspH1aVal 383
 |||||
 Db 961 GAGCTGAGAGACACCTGACCATGTT 987
 |||||

RESULT 7
 CK279064 936 bp mRNA linear EST 03-ATG-2004
 DEFINITION EST725142 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAB989 5' end, mRNA sequence.
 CK279064
 VERSION CK279064.1 GI:39836042
 KEYWORDS EST
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 936)
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST725143
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via http://genome.arizona.edu/orders/ .
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 Source
 1..936
 Location/Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAB989"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="PH10B-Tona"
 /clone_1ib="potato abiotic stress cDNA library"
 /note="Vector: PCMVSPORT6.1; Site 1: EcorI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation
 of the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	3,276-166	Length:	936
Score:	1455.00	Matches:	286
Percent Similarity:	95.51%	Conservative:	12
Best Local Similarity:	91.67%	Mismatches:	12
Query Match:	55.30%	Indels:	2
DB:	7	Gaps:	0

US-10-718-952-12 (1-510) x CK279064 (1-936)

QY 163 LeuGlnLysGlnLeuArgProTyMetGlnSerMetValProLeuProGly1leTyArgP 182
 |||||
 Db 3 CTGCAAAAGCAGCTGAGGCCCTTACATGGAATCCATGATTCCTCTCTGATCTATGAC 62
 |||||
 QY 183 ProAAspPhe1leA1aAAsnGlnGluGluArgA1aAAsnAVal1leLeuG1YThrLys 202
 |||||
 Db 63 CTGACTTCATTCATGAGCTTACCAAGACAGACACGCTGCAACAGTATCAAAAGCAAG 122
 |||||
 QY 203 GlnGlnGlnValGlnGln1leLeuAAsp1leLysA1aPheLysG1Vala1aThrLysVal 222
 |||||
 Db 123 AAAGAACAAGTTGATCAATATTTTAAAGTATTTAGAGACTTCAAGAGAAAGCAAGTA 182
 |||||
 QY 223 AAspLysValValLeu1leTyThrA1aAsnThrGluArgTYrSerAAsnLeuValValGly 242
 |||||
 Db 183 GACAAGATAGTGCTTCTATGACCTGCCAACCGAAGGTACAGCAATGTGTTGTGGC 242
 |||||
 QY 243 LeuAAspThrMetGlnAAsnLeuLeuA1aValAAspArgAAsnGlu1leSer 262
 |||||
 Db 243 CTTAACGATACCATGAAACCTTTTACTGCTGTGGATGAGAAAGAGCTGAATATCT 302
 |||||
 QY 263 ProSerThrLeuTyA1a1eA1aCyAValMetGluAAsnValProPhe1leAAsnLys 282
 |||||
 Db 303 CCTTCTACATTCGACGCTATGCTGTATCTCGAAATATGCTCTTATCATCAAGGAAC 362
 |||||
 QY 283 ProGlnAAsnThrPheValProGlyLeu1leAAspLeuA1a1eA1aAAsnThrLeu1le 302
 |||||
 Db 363 CCTCAAAACACTTTTGTTCAGGTCTTATGATTTGGCCATTAAGAGAAACCTTAATT 422
 |||||
 QY 303 GlyGlyAAspAAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAAspHeu 322
 |||||
 Db 423 GGTGTGTATGACTTTTAAAGGTGTCAACCAAGAGAAAGTCCGTGTGTGATTTCTT 482
 |||||
 QY 323 ValGlyA1aGly1leLysProThrSer1leValSerTYrAAsnH1aLeuG1YAAsnAAsp 342
 |||||
 Db 483 GTTGAAGCTGTGATTAACCAACGTCATTAAGTACCTCAATCACTTGGTAAACAATAC 542
 |||||
 QY 343 GlyMetAAsnLeuSerA1aProGlnThrPheArgSerLysGln1leSerLysSerAAsnVal 362
 |||||
 Db 543 GGAATGAATCTTTTGTCTCTCAAAACCTTCGGTCTAAGAGATCTCAAAAGTATGTT 602
 |||||
 QY 363 ValAAspAAspMetValAAsnSerAAsnA1a1leLeuTYrGluProGlyGlnH1aProAAspH1a 382
 |||||
 Db 603 GTTGTATGACATGTTGTCAGAACCTCTCTATGATCTGGGAGACACCTTACCAT 662
 |||||
 QY 383 ValValA1a1leLysTYrValProTYrValGlyAAspSerLysArgA1aMetAAspG1uTYr 402
 |||||
 Db 663 GTCTGTGATCAACTATGTTCCATATTTTGAAGACCAAGAGGCAATGATGATGAC 722
 |||||
 QY 403 ThrSerGln1lePheMetGlyGlyLysAAsnThr1leValLeuH1aAsnThrCyAAsnLys 422
 |||||
 Db 723 ATGTACAGATTTTTCATGAGGCGAAAGACCACTATGTTATGACCAACACTTGTGAGGAC 782
 |||||

QY 423 SerLeuLeuAlaAlaProIleIleLeuAspLeuValIleLeuAlaGluLeuSerThrArg 442
Db 783 TCTCTTTGGCAGCTCCAAATATCTTGAGCTTGCTCTCTCGTGAAGTCACTGCGC 842
QY 443 IlegInPheLysAlaGlu--AengLugLylsPheHisSerPheHisProValAlaThrI 462
Db 843 ATTCAAGTCAAGAGCGAAAGGGAGGGAGGAGTTCCACTCTCCACCTGTGTCTACTA 902
QY 462 leLeuSerTyrLeuThrLysAlaProLeuVal 472
Db 903 TTCTCAGCTACTCTCAACCAAGGCTCCCTGTGTA 934
RESULT 8
LOCUS CV135801
DEFINITION EST947010 Sequencing ESTs from loblolly pine embryos Pinus taeda
ACCESSION CV135801
VERSION CV135801.1 GI:51898116
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 972)
Buell,C.R., Hsiac,J., and Cairney,J.
Sequencing of ESTs from loblolly pine embryonic libraries
Unpublished (2004)
Contact: C. Robin Buell
Plant Genomics Group
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: rbuell@tigr.org
This clone is available through TIGR. Please contact pine@tigr.org
for further information.
FEATURES
Source
Location/Qualifiers
1..972
/organism="Pinus taeda"
/mol_type="mRNA"
/culturvar="7-56 mother tree, open-pollinated tree from,
lyons, Georgia, USA"
/db_xref="taxon:3352"
/clone="RP1A131"
/lab_host="E.coli DH10B-TonA"
/note="Organ: Zygotic Embryo and Megagametophyte. Somatic
/clone lib="Sequencing ESTs from loblolly pine embryos"
Embryo; Vector: pCMV-SPORT 6.1; Site_1: NotI; Site_2:
BcoRI; tissue: Whole megagametophytes isolated from pine
seeds, whole embryos excised from these megagametophytes,
whole somatic embryos and suspensor tissue from tissue
culture, isolated from cell line A12. Pooled RNA from
zygotic embryos, megagametophytes, and somatic embryos was
used for library construction. Pine cones were harvested
weekly from open-pollinated 7-56 mother trees, collections
occurred from 7/01/02 until 10/15/02. Whole
megagametophytes were first isolated from pine seeds, and
whole embryos excised from these megagametophytes each was
flash frozen. Embryo development was assessed using the
system of Pullman et al (Pullman GS, Johnson S, Peter G,
Cairney J, Xu N. 2003. Loblolly pine somatic
embryogenesis: development of a maturation medium and
resulting embryo quality. Plant Cell Reports 21:747-758
(http://link.springer.de/link/service/journals/00299/conte
nts/03/00586/). For photographs see Clavatta et al 2001.
Clavatta VT, Morillon R, Pullman GS, Christeels M,
Cairney J. 2001. An aquaglyceroporin is abundantly
expressed early in the development of the suspensor and
the embryo proper of loblolly pine (Pinus taeda L.). Plant
Physiol. 127: 1556-1567
(http://www.plantphysiol.org/cgi/content/full/127/4/1556) ."

Alignment Scores:

Pred. No.: 4,12e-164 Length: 972
Score: 1438.00 Matches: 284
Percent Similarity: 95.05% Conservative: 23
Best Local Similarity: 87.93% Mismatches: 16
Query Match: 54.66% Indels: 2
DB: 7 Gaps: 0

US-10-718-952-12 (1-510) x CV135801 (1-972)

QY 183 ProSerPheIleAlaAlaAsnGlnGluAspValIleLeuValIleLysGlyThrLys 202
Db 1 CCGGATTCATAGCGCGCAACCAAGGTGAGGGCGCAATTAATGATCAAGGCTCCAG 60
QY 203 GlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThrLysVal 222
Db 61 AAAGAGCACTGCAAAAGATTTTCAGAGTTTCAGAGTTTAAATGAAACAAAGGTG 120
QY 223 AspLysValValIleuTyrThrAlaAsnThrGluArgTyrSerAsnLeuValValGly 242
Db 121 GACAAGTAGTCGACTGTGAGTGTGATACCGAAGCATACAGCGAGCTATCGTGGGT 180
QY 243 LeuAsnAspThrMetGluAsnLeuLeuAlaAlaValAspArgAsnGluAlaGlnIleSer 262
Db 181 TTAATATGACAGAAAGAAACCTCTGCTCTGGAAGAAATGAGTCGAAATCTCG 240
QY 263 ProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySer 282
Db 241 CCATCGACTCTCTACGGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 283 ProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIle 302
Db 301 CCTCAAAACACCTTCGTTCCAGATTGATCATTTAGCAATTAACAAAGACAGTATCATC 360
QY 303 GlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValIleValAspPheLeu 322
Db 361 GCAGAGATGACTTCAAGAGCGGCGAGCAAGATGAACTGCTGCTGCTGCTGCTGCTGCT 420
QY 323 ValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAspAsnAsp 342
Db 421 GTCCGTGCGGCTCTTAAGCCAACTTCGATTTGAGCTACCAATTCCTCGTAACAAGCAG 480
QY 343 GlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSerAsnVal 362
Db 481 GGATGATCATCTGAGCTCCCAAGACCTTCGCGCAAGAAATTTCAAGAGCAATGTC 540
QY 363 ValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisProAspHis 382
Db 541 GTCCAGCATAGTTGCCAGCAAGCCCATTTTACGAACCAAGGAGAACCCAGACCAT 600
QY 383 ValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAspGluTyr 402
Db 601 GTCTGCTCATCAAGTATGTACCTTACGTTGGAACAGCAAGAGACCATGAGAGTAC 660
QY 403 ThrSerGluIlePheMetGlyLysAsnThrIleValLeuHisAsnThrCysGluAsp 422
Db 661 ACGTCCAGATATTATTTAGGAGGAAACACCTTGTCATCCACATATCATGGAAGAT 720
QY 423 SerLeuLeuAlaAlaProIleIleLeuAspLeuValIleLeuAlaGluLeuSerThrArg 442
Db 721 TCGCTCTTCTTGAGCTCCCTCATCTTGTAGCTGTGTGCTGCGCGAAGCTGTACAGG 780
QY 443 IlegInPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIle 462
Db 781 ATTCACTTAAGAAAGAAAGC-GAGGGGAATTTCAATCTTTCCACCGGTAGCAAGCTT 839
QY 463 LeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsnAlaLeu 482
Db 840 TTAAGTCACTCAACCAAGGCGCACTGTGTACCAACAGGCAAGCTGTGTAAAGCCCTG 899
QY 483 SerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGlu 502
Db 900 CGGAAGCAGAGGCGAGTGTGAGAC-ATCATGAGGGGCTGCTGCGCTGCTGCCGAA 958

QY 503 AspaanMet 505
Db 959 AACCAACATG 967

RESULT 9
CRS99696
LOCUS
DEFINITION CRS99696 1613 bp mRNA linear HTC 21-JUN-2004
full-length cDNA clone CS0D1009YP03 of Placenta Cot 25-normalized
of Homo sapiens (human).

ACCESSION CRS99696
VERSION CRS99696.1 GI:50480503
KEYWORDS HTC: CNSLT cDNA.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1613)
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1613)

REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. 1613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1009YP03"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:
Pred. No.: 1,2e-163 Length: 1613
Score: 1437.00 Matches: 276
Percent Similarity: 70.75% Conservative: 82
Best Local Similarity: 54.55% Mismatches: 88
Query Match: 54.62% Indels: 60
Gaps: 4

US-10-718-952-12 (1-510) x CRS99696 (1-1613)

QY 6 PheylValAlGusEPrObanValYsYrThrGluThrGluIleGlnSerValYrAsn 25
Db 56 TTCCTGCGAGAGCCGAGCGTGTCTACGGCCCGAGGCGATGAGCGCAATACGAG 115
QY 26 TyrGluThrGluLeuValHisGluAenAArganGlyThrTyGluTrpIleValYs 45
Db 116 TACCGAGCAGCCGCGTCAGCCGCAAG----- 142
QY 46 ProLysSerValAsnTyGlnPheLysThrAsnThrHisValProLysLeuGlyValMet 65
Db 142 ----- 142
QY 66 LeuValIGlyTrpGlyGlyAsnAsnGlySerThrLeuThrGlyValIleAlaAsnArg 85
Db 143 -----GGTGGCGTT----- 151
QY 86 GluApeIleSerTrpAlaThrLysApeLysIleGlnGlnAlaAsnTyPheGlySerLeu 105
Db -----

Db 152 -----CTCAAGAGGCCAAGCTACTAGGCTGCTG 181
QY 106 ThrGlnAlaSerAlaIleArgValGly--SerPheGlnGlyGluGluIleTyAlaPro 124
Db 182 ACTCAGCGCGGCGACCGTGAAGCTGGCGCTGACGCGCGAGGCGCAGAGAGTGTTCCTACCC 241
QY 125 PheLysSerLeuLeuProMetValAsnProAspApeIleValPheGlyGlyTrpApeIle 144
Db 242 TTCACCGCGGCTGTCCCATGTGGTGGCCCAACGACTGTGTGTGATGTGGTGGACATC 301
QY 145 SerAenMetAenLeuAlaAspAlaMetAlaArgAlaValPheAspIleAspLeuGln 164
Db 302 TCGTCGCGGAACCTGCGCGAGCGCATGCGCGCGCGGAGGTCTGTGACTGGCGGCTGCGAG 361
QY 165 LysGlnLeuAArgProTyrMetGluSerMetValProLeuProGlyIleTyAspProAsp 184
Db 362 GAGCAACTGTGGCCCGCATGTGAGGCGCTCGCGCGCCCGCTTCGTATTACATCCCCGAA 421
QY 165 PheIleAlaIleAsnGlnGluGluArgAlaAsnAsnValIleValGlyThrLysGlnGlu 204
Db 422 TTCATGCGGCGCCACCAAGCGCGCGCGGACAACTCATGCCAGGCTCGCGTCCGAG 481
QY 205 GlnValGlnGlnIleIleLysApeIleLysApeLysAlaPheLysGluAlaThrLysValAspYs 224
Db 482 CAGCTGAGCAGAGTCCGAGGAGGACATCCGAGACTTCGCGTACGCGCGGCGCTGACAAA 541
QY 225 ValValValLeuTrpThrAlaAsnThrGluArgTySerAenLeuValGlyLeuAsn 244
Db 542 GTCATAGTGTGTGGACGCGCAACGAGCGCTTCGTGAGTGATTCAGGCTCAAC 601
QY 245 AspThrMetGluAenLeuLeuAlaValAspAArganGluAlaGluIleSerProSer 264
Db 602 GACACAGCGCGAAGACTCTCTCGCACCATGTGACTCGGT--CTGAGAGTGTGCGCTCC 658
QY 265 ThrLeuTyAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySerProGln 284
Db 659 AGCTCTTCGCGCGTGGCCAGCATCTTGAGAGGCTGTGCTTCTCAATGCGTCCGAG 718
QY 285 AenThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIleGlyGly 304
Db 719 AACACCTGTGTGCGCGAGCTCTTGAGCTCGGTGCGAGCACCGGGTTTGTGTGGCGGA 778
QY 305 AspApePheLysSerGlyGlnThrLysMetLysSerValIleValAspPheLeuValGly 324
Db 779 GATGACTTCAGTCAAGTCAAGCGCAGACCAAGTCAAGTCCGCTGTGTGACTTCTCAATTGAC 838
QY 325 AlaGlyIleLysProThrSerIleValSerTyrAenHisLeuGlyAsnAspApeIleMet 344
Db 839 TCCGCGCTCAAGACCATGTCCATCTGTGATTCAACCACTGCGGCAACAGATGGGAG 898
QY 345 AenLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSerAsnValAlaAsp 364
Db 899 AACCTATCGCGCCATTGACGTTCCGCTCAAGAGAGGTGCCAAGACAGCATGGTGGAC 958
QY 365 AspMetValAsnSerAenAlaIleLeuTyTrpGluProGlyGluHisProAspHisValVal 384
Db 959 GACATGTGTGACAGAGCAACCAAGTCTTATACGCCCGCGCAAGACCTGACACATGCGCTG 1018
QY 385 ValIleLysTyValAlaProTyrValGlyAspSerLysAArgAlaMetAspGlyTyThrSer 404
Db 1019 GTCATCAAGTATGTGCTGTACGTGTGACAGCAAGCGCGCGTGTGATGATTAATCTTCG 1078
QY 405 GlnIlePheMetGlyGlyLysAsnThrIleValIleuHisAenThrCysGluAspSerLeu 424
Db 1079 GAGCTGATGCTGGCGGAGCAACACACTGTGTGCAACAACGTGAGAGACTGCTG 1138
QY 425 LeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGln 444
Db 1139 CTGGCGGACCATCATCATGTGAGCTGTGACCGGCTGTGACCGGCTGTGACCGGCTGTGAGC 1198
QY 445 PheLysAlaGluAenGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSer 464
Db 1199 TTCGTCACTGACATGAGCCGAGCCGAGCCTTCCACCCGCTGTGTCTCTGCTGACG 1258

QY 465 TyrleuThrlYsAlaProleuValProProGlyThrProValValAsnAlaIleuSerlys 484
:::|||||
Db 1259 TTCTCTTCAGAGGGCCCACTAGTGCCTGCCGCCGACAGCCCGTGTGTCATAGCCGTTTTCGCCG 1318
QY 485 GlnATgAlaMetleuGluAsn11eMetArGa1aCyValAGlyLeuAlaProGluAsnAsn 504
:::|||||
Db 1319 CAGCGCAGCTGCTGAGAAACACTCTCTCAGGGCTGCGTGGGCTCCGCCACAGAACAC 1378
QY 505 MetIleuGluTyrTlys 510
:::|||||
Db 1379 ATGCTCTGGAACACAA 1396

RESULT 10
CK275352 914 bp mRNA linear EST 03-AUG-2004
LOCUS CK275352
DEFINITION EST721430 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POADM85 5' end, mRNA sequence.
ACCESSION CK275352
VERSION CK275352.1 GI:39832330
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 914)
Buell, C.R., Hart, A., Zismann, V., Karameyheva, S.A. and Baker, B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
CONTACT: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TNG.
Location/Qualifiers
1..914
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POADM85"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: pCMVSPORT1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN
Alignment Scores:

Pred. No.: 8,26e-163 Length: 914
Score: 1427.00 Matches: 277
Percent Similarity: 97.95% Conservative: 9
Best Local Similarity: 94.86% Mismatches: 6
Query Match: 54.24% Indels: 0
DB: 7 Gaps: 0

US-10-718-952-12 (1-510) x CK275352 (1-914)

QY 217 LysGluAtrThrlYsValAspLyValValIleuThrAlaAsnThrGluArTyr 236
:::|||||
Db 1 AAGAGAAACAGCAAGTGAAGCAAGTGTGTTTGTGAGCTGCCAACAACAGAAATAC 60
QY 237 SerAsnLeuValValAGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArg 256
:::|||||
Db 61 AGCAGTGTGCTGTGGCTTAATGATACCATGAGAAACCTCTTGGCTTGTGATAGA 120
QY 257 AsnGluAglIleSerProSerThrLeuThrAlaIleAlaCyValMetGluAsnVal 276
:::|||||
Db 121 AATGAGGCCGAATAATCTCTCCACCTGTATGCTATTGCTGTATTTGAAATAATGTG 180
QY 277 ProbheilleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIle 296
:::|||||
Db 181 CCTTTCATCAATGGAAGCCCAACAACATTTCCTCCAGGCTCATGATTTGCCATC 240
QY 297 AlaArgAsnThrlleuIleGlyIleAspAspPheIysSerGlyInThrIysMetIysSer 316
:::|||||
Db 241 AAGAGAAACCTTGTGATGGTGTGATGACTTTAAGAGTGTGCAACCAAGATGAGTCA 300
QY 317 ValIleuValAspPheLeuValAGlyAlaGlyIleIleYsProThrSerIleValSerTyrAsn 336
:::|||||
Db 301 GTGCTGTGTATTTCTCTTGTGGAGCTGTATTAAGCAACATAAATGTAGCTACAAAC 360
QY 337 HisleuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIysGlu 356
:::|||||
Db 361 CATTTGGTAAATGATGATGAATGAATATCTGCGGCCCTCAGACTTTCGCTCAAGAG 420
QY 357 IleSerIysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluPro 376
:::|||||
Db 421 ATCTGAAACAAATGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 377 GlyGluHisProAspPheValValIleIleIleYsValProTyrValAGlyAspSerIys 396
:::|||||
Db 481 GGAAGAGCACTCTGACCAATGTTGTATTAAGATATCAATATGATGATGATGATGATGAT 540
QY 397 ArgAlaMetAspGluTyrThrSerGluIlePheMetGlyIysAsnThrlleValIleu 416
:::|||||
Db 541 AGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 417 HisAsnThrCysGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleu 436
:::|||||
Db 601 CACAATACCTGTGAGATTTCTTTTGGCTGCTCAATATATCTTGATGATGATGATGATGAT 660
QY 437 AlaGluIleuSerThrArgIleGlnPheIysAlaGluAsnGluGlyIysPheHisSerPhe 456
:::|||||
Db 661 GCTAATCTACACACCCGATTCACCTCAAGCTAAGAGAGGGATGATGATGATGATGATGAT 720
QY 457 HisProValAlaThrIleLeuSerIysLeuThrIysAlaProLeuValProProGlyThr 476
:::|||||
Db 721 CATCCGTGGGAGCAATCTCTAGCTATCTTACCAAGGCTCCTCTGTGATCCACAGATACA 780
QY 477 ProValValAsnAlaLeuSerIysGlnArgAlaMetleuGluAsnIleMetArgAlaCys 496
:::|||||
Db 781 CCAAGTGTGATGCTCTTTCAACAGAGGGCATGCTTGAGAACATATATGAGGGCTTGT 840
QY 497 ValGlyLeuAlaProGluAsnMetIleleuGlu 508
:::|||||
Db 841 GTTGAGCTGGCACCAAGAACATCATGATATTGGAA 876

RESULT 11
CO414034 953 bp mRNA linear EST 02-JUL-2004
LOCUS CO414034
DEFINITION EST844419 Sequencing ESTs from loblolly pine embryos Pinus taeda

CDNA clone P1AM182 5' end, mRNA sequence.
 ACCESSION COA14034
 VERSION COA14034.1 GI:49630282
 KEYWORDS EST.
 SOURCE Pinus taeda (loblolly pine)
 ORGANISM Pinus taeda
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
 1 (bases 1 to 953)
 REFERENCE Buell, C.R., Hsiao, J. and Cairney, J.
 TITLE Sequencing of ESTs from loblolly pine embryonic libraries
 JOURNAL Unpublished (2004)
 CONTACT: C. Robin Buell
 COMMENT Plant Genomes Group
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: rbuell@tigr.org
 This clone is available through TIGR. Please contact pine@tigr.org for further information
 Seq primer: APT TAG GTG ACA CTA TAG.

FEATURES

Source

1..953
 /organism="Pinus taeda"
 /mol_type="mRNA"
 /cultivar="7-56 mother tree, open-pollinated tree from, Lyons, Georgia, USA"
 /db_xref="taxon:3352"
 /clone="P1AM182"
 /lab_host="E.coli DH10B-TonA"
 /clone_lib="Sequencing ESTs from loblolly pine embryos"
 /note="Organ: Zygotic Embryo and Megagametophyte, Somatic Embryo; Vector: pCMV-SPORT 6.1; Site_1: NotI; Site_2: EcoRV; tissue: Whole megagametophytes isolated from pine seeds, whole embryos excised from these megagametophytes, whole somatic embryos and suspensor tissue from tissue culture, isolated from cell line A12. Pooled RNA from zygotic embryos, megagametophytes, and somatic embryos was used for library construction. Pine cones were harvested weekly from open-pollinated 7-56 mother trees, collections occurred from 7/01/02 until 10/15/02. Whole megagametophytes were first isolated from pine seeds, and whole embryos excised from these megagametophytes each was flash frozen. Embryo development was assessed using the system of Pullman et al (Pullman GS, Johnson S, Peter G, Cairney J, Xu N. 2003. Loblolly pine somatic embryogenesis: development of a maturation medium and resulting embryo quality. Plant Cell Reports 21:747-758 (http://link.springer.de/link/service/journals/00299/contents/03/00586/). For photographs see Ciavatta et al 2001. (Ciavatta VT, Morillon R, Pullman GS, Chrispeels M, Cairney J. 2001. An aquaglyceroporin is abundantly expressed early in the development of the suspensor and the embryo proper of loblolly pine (Pinus taeda L.). Plant Physiol. 127: 1556-1567 (http://www.plantphysiol.org/cgi/content/full/127/4/1556))".

ORIGIN

Alignment Scores:
 Pred. No.: 3,59e-162 Length: 953
 Score: 1422.00 Matches: 273
 Percent Similarity: 94.62% Conservative: 26
 Best Local Similarity: 86.35% Mismatches: 17
 Query Match: 54.05% Indels: 1
 DB: 7 Gaps: 0

US-10-718-952-12 (1-510) x COA14034 (1-953)

QY 115 SerPheGlnGlyGluIleTyrAlaProPheIysSerLeuLeuProMetValAsnPro 134
 Db 7 TCCTTTACGGCGAGAGATTCATCTCTTTCAAGACCTTCTCCATGCGATGATCT 66
 QY 135 AspAspIleValPheGlyGlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetVal 154

Db 67 AATGAATGCTCTTTGGTGGATGGACATTAACAACATGAATCTGCGGATCCATGGCC 126
 QY 155 ArgAlaIysValPheAspIleAspLeuGlnIysGlnLeuAAGProTyrMetGluSerMet 174
 Db 127 CGAGCCAGAGGTTCTTGACATTCGACCTTAAGAGAGCTGGCGCTTACATGCAAGGCATG 186
 QY 175 ValProLeuProGlyIleTyrAspProAspPheIleAlaIAsnGlnGluIuArgAla 194
 Db 187 ACTCTCTTCTCGGTATATATGACCCGGAATTCATAGCGGCACCAACAGGTGAGGGCA 246
 QY 195 AsnAsnValIleIysGlyThrIysGlnGlnIuValGlnGlnIleIleIysAspIleIys 214
 Db 247 AATATGTCATCAAAAGCTCCAGAAAGCAACGCAAAAGATTAATCAAGATATCA 306
 QY 215 AlaPheIysGluAlaThrIysValAspIysValValIleTyrThrAlaAsnThrGlu 234
 Db 307 GATTTTAATCGAAMAACAAGGTGACAAAGTAGCTGATCTGTGACTGTAATACCGAA 366
 QY 235 ArgTyrSerAsnLeuValValGlyIleuAsnAspThrMetGluAsnLeuAlaIalVal 254
 Db 367 CGATACAGCGAGTTATGCTGGGTTTAATGACAGAAAGAAACCTTACTGCTTCTTG 426
 QY 255 AspArgAsnGluAlaGluIleSerProSerThrIleuTyrAlaIleAlaCysValMetGlu 274
 Db 427 GAAGAATGATGATCGGAATCTCGGCATCGACTCTCTAGCGGTTGGCTTGACAGAG 486
 QY 275 AsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeu 294
 Db 487 AACGTTCCCTTCATTAACGGCAGCTCTCAAAACCTTCGTTCCAGGATTCATTTA 546
 QY 295 AlaIleAlaAsnThrLeuIleGlyIysAspAspPheIysSerGlyIleThrIysMet 314
 Db 547 GCAATACAAAGAACAGATATCGACAGAGATGACTTCAAGAGCGGCGACAAAGATG 606
 QY 315 LysSerValLeuValAspPheLeuValGlyIleIysProThrSerIleValSer 334
 Db 607 AAGTCGACTCTCGGATTTCTTAGTCGCTCCGCTTAAACCACTTCATTTGTAGC 666
 QY 335 TyrAsnHisLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSer 354
 Db 667 TACATCATCTCGGTAACACGAGCGGATGATGTGACGCTCCCAACATTCGCGTCA 726
 QY 355 LysGlnIleSerIysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyr 374
 Db 727 AAGGAAATTTAAAGAGCAAGTGTGTCAACAGATGTTCAGCAAGCCATCTTTAC 786
 QY 375 GluProGlyGluHisProAspHisValValIleIysTyrValProTyrValGlyAsp 394
 Db 787 GACCAAGGGAGAACCCAGACCATGTCTCTCATCAAGTATGTAATCTTACGTTGAGAC 846
 QY 395 SerIysArgAlaMetAspGluTyrThrSerGluIlePheMetGlyIysAsnThrIle 414
 Db 847 AGCAAGAGAGCCATGAGACGATGACACGTCGAGATATTATGGAGGAAA-AAACCCCTT 905
 QY 415 ValLeuHisAsnThrCysGluAspSerLeuLeuAlaIalProIleIle 430
 Db 906 GTAATCCACAATATACATGGAAGATTGCTTGTGACGCTCTGTATC 953

RESULT 12
 LOCUS CO113870
 DEFINITION GR_EB014121.r GR_Eb Gossypium raimondii cDNA clone GR_EB014121
 3', mRNA sequence.
 ACCESSION CO113870
 VERSION CO113870.1 GI:48812557
 KEYWORDS EST.
 SOURCE Gossypium raimondii
 ORGANISM Gossypium raimondii
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 898)

AUTHORS

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
Wing, R.A.

TITLE

Global assembly of Cotton ESTs

JOURNAL

Unpublished (2004)

COMMENT

Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 014 row: 1 column: 21.

FEATURES

source

Location/Qualifiers

1..898

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_EB014121"

/tissue_type="floral"

/dev_stage="3 to +3 DPA"

/lab_host="DH10B"

/clone_id="GR_EB"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:

EcoRV; library made by Invitrogen with RNA supplied by

Wendle lab. Directional cloned into NotI-EV. Clones

plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.: 1.99e-158 Length: 898
Score: 1391.00 Matches: 263
Percent Similarity: 95.65% Conservative: 23
Best Local Similarity: 87.96% Mismatches: 13
Query Match: 52.87% Indels: 0
DB: Gaps: 0

US-10-718-952-12 (1-510) x CO113870 (1-898)

49 ValAenTYrGlnPheIySThrAsnThrHISValProLYSLeuGLYValMetLeuValGLY 68
Db 1 GTCAATATGAAATTCAAGACATGATACCATGCCCTTAATTGGGGTATGCTTGGA 60
Qy 69 TTPGLIYAsnAsnGLYSerThrLeuThrGLYValIleAlaAsnArgGLUAspIle 88
Db 61 TGGGAGAGAAACAATGTTCAACCTCACCGGTGTTATAGCTAACAAAGAGGATATC 120
Qy 89 SerTPAATnTrLyAspLyValIleGlnGlnAlaAsnTYrPheGLYSerLeuThrGlnAla 108
Db 121 TCTTGGGTACTTAAGACAGAGTCAACAGCGCTAATTGCTTGCATTGACTCAAGCA 180
Qy 109 SerAlaIleArgValISerPheGlnGLYGLUValIleTYrAlaProPheLYSerLeu 128
Db 181 TCACAGATCCGAAATGGCTTTCACATGAGAGAGATTATGCTCCATTAAAGTCTT 240
Qy 129 LeuProMetValAsnProAspAspIleValPheGLYTYrPASPILeserAsnMetAsn 148
Db 241 CTTCTTAATGAGTAAACCAATGATATGTGTTTGAAGATGGACATTAGTACATGAAC 300
Qy 149 LeuAlaAspAlaMetAlaArgAlaIleValPheAspIleAspLeuGlnIleuArg 168
Db 301 CTAGCTGATGCAATGGCTTAGGGCCAAAGTTTTCGATCGATGATCTGCAAAAGCACTGAGA 360
Qy 169 ProTYrMetGluSerMetValProLeuProGlyIleTYrAspProAspPheIleAlaIle 188
Db 361 CCTTCAATGAGATCCATGCTCCCTCGAATCTHAGATCTGATTTCTATCTCTCT 420
Qy 189 AsnGlnGLUArgAlaAsnAsnValIleLYSGLYThrLYSGlnGLUValGlnGln 208
Db 421 AACCAAGGTGAACGTGCCAATATGTCATCAAGGGAGCAAGAAAGAACAGTTCAAGCAG 480
Qy 209 IleIleLYSAspIleLYSAlaPheLYSGlnValAlaTrLYSValAspLYSValValLeu 228
:::|||||

Db 481 GTCAATCAAGACATCAAGAGACTCAAGAGAAAAACAAGGTGACAGAGTTGTACTC 540
Qy 229 TTPThAlaAsnThrGLUArgTYrSerAsnLeuValGLYLeuAsnAspThrMetGln 248
Db 541 TGAATGCAAAACATGAGAGAGTACAGCAATGTCAATCGGGCTTAATGACACCGTGGA 600
Qy 249 AsnLeuAlaAlaValAspArgAsnGLUAlaGlnIleSerProSerThrLeuTYrAla 268
Db 601 AGCCTTAGGCTTTTGGAGAAAGATGAATCAAGATTCTCTTCACATTTATAGCT 660
Qy 269 IleAlaCysValMetGluAsnValProPheIleAsnGLYSerProGlnAsnThrPheVal 288
Db 661 ATGCTTGTTGTTCTTGAATGTTCTCTTCTTCATAGTGCACCCACAAAACCTTGTCT 720
Qy 289 ProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIleGLYAspAspPheLYs 308
Db 721 CCAGGGTGAATTGATTGGCTATTCAAGAGAACTGTCTGATTGAGAGATGACTTCAG 780
Qy 309 SerGLYnTrLYSMeLYSserValLeuValAspPheLeuValGLYAlaGLYIleLYs 328
Db 781 AGTGGCCAGACCAAGATGANAATCTGCTCTCTGATTCTCTTGTGGGGCTGGATCAAG 840
Qy 329 ProThrSerIleValSerTYrAsnHISLeuGLYAsnAsnAspGLYMetAsnLeuSer 347
Db 841 CCAACATGATGATGAGATTACCACTGAGAAATATATGATGACATGATCTGTCA 897

RESULT 13
LOCUS CO085839 865 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_Ea03E19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea03E19.5',
mRNA sequence.

ACCESSION CO085839
VERSION CO085839.1 GI:48764473
KEYWORDS EST.

SOURCE
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Malvales; Malvaceae; Malvaceae; Gossypium.

REFERENCE
1 (bases 1 to 865)
Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
Wing, R.A.

TITLE
JOURNAL
COMMENT
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 03 row: E column: 19.

FEATURES
source
Location/Qualifiers
1..865
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea03E19"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_id="GR_Ea"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clones
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.: 1.69e-156 Length: 865
Score: 1375.00 Matches: 264
Percent Similarity: 97.22% Conservative: 16


```

Best Local Similarity: 91.67%      Mismatches: 8
Query Match: 52.26%      Indels: 0
DB: 7      Gaps: 0
US-10-718-952-12 (1-510) x CD085839 (1-865)

QY 143 AapIleSeSaenMeAenLeuAlaAapAlaMeAlaAaArgAlaLeValPheAapIleAap 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 GACATTAGGACATGAACTTACCTGATCTTATGGCTAGGAGCCCAAGTTTCCAGCATGAT 61

QY 163 LeuGlnLySgInLeuArgProTyMeGusSerMeValProLeuProGlyIleTyAap 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 CTGCAAAAGCACTGAGACCTACATGATGATCATGATGATCCATCCCTCGAATCTACAT 121

QY 183 ProAapPheIleAlaAlaAenGlnGluGAlaAenAenValIleYsgIyThrLyS 202
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 CCGATTTCATGCTGCTCAACCAAGGTGAACGCGCAATTAATGTCATCAAGGAGCAAG 181

QY 203 GlnGlnGlnValGlnGlnIleIleYsAapIleYsAlaPheLySgInAlaThrLySVal 222
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 AAGAAGCAAGTTCAAGAGGATCATCAAGACATCAAGAGTTCAAGAGAAACCAAGGTG 241

QY 223 AapLySValValIleuTrpThraAaenThrgUaArgTySerAenLeuValValGly 242
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 GACAAAGTTGTTGTTCTGAGCTGCAACCACTGAGAGGTAACGATCTCATCGGGG 301

QY 243 LeuAenAapThMeGluAenLeuLeuAlaAlaValAapArgAenGluAglIleSer 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 CTAAATGACACCGGTGAAGCCTTATGCTCTTGGAGAGATGAATCAGAGATTCT 361

QY 263 ProSerThLeuTyAlaIleAlaCySValMeGluAenValProPheIleAenGlySer 282
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 CTTTCACTTTGATGCTATGCTGCTGCTTGAAGAGTTCTTTCATCATAGGACAC 421

QY 283 ProGlnAenThPheValProGlyLeuIleAapLeuAlaIleAlaArgAenThrLeuIle 302
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 422 CCACAAACACCTTGTCTCAGAGGTTGATTTGATTTGGCTATTCAAGAGATGTCGAT 481

QY 303 GlnGlyAapAapPheLySergIyGlnThrLySMeLySergValLeuValAapPheLeu 322
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 GGAGAGAGTACTTCAAGAGTGGCCGACCAAGATGAATCTGCTCTCGTGAATTTCTT 541

QY 323 ValGlyAlaGlyIleYsProThrSerIleValSerTyAaenHileuGlnYsAenAap 342
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 GTTGGGCTGGGATCAAGCAACATGATGATGATTAACCAACATTTGGAATTAATGAT 601

QY 343 GlnMeAenLeuSerAlaProGlnThrPheArgSerLySgInIleSerTySerAenVal 362
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 602 GGCATGAATCTGTCAAGCAACCAACCTTCCTCCATCAAGAGATCTCAAGAGCATGTT 661

QY 363 ValAapAapMeValAenSerAenAlaIleuTyArgIleProGlyGlnHileProAapHis 382
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 662 GTTGTATGATGCTTCAAGCAATGGAATCCGTATGAGCCCTGAGTAATCTGATCAT 721

QY 383 ValIleValIleIleTyArgValProTyArgValIleYsAenSerLySgAlaMeLySgIleTy 402
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 722 GTTGTGTATGATCAAGATGATGCTATGATGAGAGCAAGAGCCCTGATGATGATC 781

QY 403 ThrSerGlnIlePheMeGlyLySgAenThrIleValIleuHileAenThrCySgIleAap 422
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 782 ACATCAAGATATTTCAATGGAGGCAAGAACCATTTGTCTTCAACACATGTAGAT 841

QY 423 SerLeuLeuAlaAlaProIleIle 430
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 842 TCCCTGTTGGCTGCTCCCATATATC 865

RESUL14
CD438045 883 bp mRNA linear EST 03-JUN-2003
LOCUS EL01N0508B11.b Endospem_5 Zea mays cDNA, mRNA sequence.
DEFINITION CD438045
ACCESSION CD438045.1 GI:31353688
KEYWORDS EST.

```

```

SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 883)
AUTHORS Lai,J., Dey,N., Kim,C.S., Bharti,A.K., Rudd,S., Mayer,K.F.,
TITLE Characterization of the maize endospem transcriptome and its
JOURNAL comparison to the rice genome
COMMENT Genome Res. 14 (10), 1932-1937 (2004)
CONTACT: Lai, Jinseng
Dr. Joachim Messing's lab
Wakeman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@wakeman.rutgers.edu
Seq primer: 73.

FEATURES
source location/Qualifiers
1..883
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/issue_type="Endospem of 7-23DAP"
/clone_id="Endospem 5"
/notes="Vector: pBluescript SK-; Site_1: EcorI; Site_2:
XhoI"

ORIGIN
Alignment Scores: 2,316-156 Length: 883
Pred. No.: 1374.00 Matches: 266
Score: 95.19% Conservative: 11
Percent Similarity: 91.41% Mismatches: 14
Best Local Similarity: 52.22% Indels: 0
Query Match: 6 Gaps: 0
DB: 0

US-10-718-952-12 (1-510) x CD438045 (1-883)

QY 143 AapIleSeSaenMeAenLeuAlaAapAlaMeAlaAaArgAlaLeValPheAapIleAap 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 GACATTAGGACATGAACTTACCTGATCTTATGGCTAGGAGCCCAAGTTTCCAGCATGAT 70

QY 163 LeuGlnLySgInLeuArgProTyMeGusSerMeValProLeuProGlyIleTyAap 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 CTGCAAGAGCACTGAGACCTACATGAGTCCATGTCATGTCCTCCGATATCATGAT 130

QY 183 ProAapPheIleAlaAlaAenGlnGluGAlaAenAenValIleYsgIyThrLyS 202
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 CCGGACTTCAATCGCGCTCAACCAAGGCTCTCGCGCAACAGTGTATCAAGGCAAG 190

QY 191 AAGGACAGGTGAGACCAATCATCAAGATATCAAGAGGCTTAAAGAGAAACCAAGT 250
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 250 AAGGACAGGTGAGACCAATCATCAAGATATCAAGAGGCTTAAAGAGAAACCAAGT 250

QY 223 AapLySValValIleuTrpThraAaenThrgUaArgTySerAenLeuValValGly 242
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 GACAAAGATAGTTGTTCTGAGCTGCAACCACTGAAAGGTATAGCATGTGCGCTGAT 310

QY 243 LeuAenAapThMeGluAenLeuLeuAlaAlaValAapArgAenGluAglIleSer 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 CTCAACGACAGATGAGAAATCTACTGCACTGTGGAACAAGACGAGCGGAGGTATCA 370

QY 263 ProSerThLeuTyAlaIleAlaCySValMeGluAenValProPheIleAenGlySer 282
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 263 CCATCAACCTATATGCAATTTGCTGTCTCATGAGGAGGAGGCTTATCATATGAGAGC 430

QY 283 ProGlnAenThPheValProGlyLeuIleAapLeuAlaIleAlaArgAenThrLeuIle 302
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 CCCGAGACACCTTGTGCTGAGGCTGATGATTTGCTATTAATAAACAACATGCTTGAAT 490

```

QY 303 GlycylAsparaphenylSerylGlyInThrIlyMetIlySerValIleuValAspHeu 322
Db 491 GGTGTGACGACTTCAAGAGTGACAGACCAAGAAATGCTGTCGATTCCTT 550
QY 323 ValGlyAlaGlyIleIyProThrSeriIleValSeriYAsnHisLeuGlyAsnAsnAsp 342
Db 551 GTTGTGCTGGAATTAAGCCCACTCAATCGAGCTTACACCACTTGGAAACAACAT 610
QY 343 GlyMetAsnLeuSeriAlaProGlnThrPheArgSeriIyGluIleSeriIySerAsnVal 362
Db 611 GGCATGAACCTGTCTGCCCTCAACATTCAGGTCCAAAGAGATCTCCAAAGCAACGTG 670
QY 363 ValAspAspMetValAsnSeriAsnAlaIleuIyTrGluProGlyGluHisProAspHis 382
Db 671 GTGGATGACATGCTTCAGACAAATGCCATCCTCTGTAGCCCGGACCATCCCATCAT 730
QY 383 ValValAlaIleIyIyTrValProIyTrValGlyAspSeriIyAspAlaMetAspGluTr 402
Db 731 GTCTGTTCATCAATATATGTCCTGATCGTGGAGACAGACAGGCGCTATGACAGATAC 790
QY 403 ThrSeriGluIlePheMetGlyIyIyAsnThrIleValIleuHisAsnThrCysGluAsp 422
Db 791 ACCTGAGAGATCTTATGCGGCGCAAGAACCATCGTCTGCACAACACCTGTGAAGAC 850
QY 423 SerIleuLeuAlaAlaProIleIleIyAspLeu 433
Db 851 TCGCTCTCTCGCGCACCTATCATCTTGATCTG 883

RESULT 15
CV290142 891 bp mRNA linear EST 23-SEP-2004
LOCUS acf01-10m3-d10 Acf01 Asparagus officinalis cDNA clone
DEFINITION acf01-10m3-d10 5', mRNA sequence.
ACCESSION CV290142
VERSION CV290142.1 GI:52575144
KEYWORDS EST.

SOURCE Asparagus officinalis (garden asparagus)
ORGANISM Asparagus officinalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
Asparagaceae; Asparagus.
1 (bases 1 to 891)
dePamphilis,C., Carlson,J., Ma,H., Soltis,D., Soltis,P.,
Opentheimer,D., Frolich,M., Doyle,J., Tanksley,S., Webb,M.,
Leebens-Mack,J., Landherr,L., Ilut,D., and Wall,K.
Generation of ESTs from early male inflorescences of Asparagus
officinalis
Unpublished (2004)

JOURNAL
COMMENT Contact: Claude dePamphilis or James Leebens-Mack
Mueller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6433
Fax: 814 865 9131
Email: cwj3@psu.edu or jh110@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (<http://pgn.cornell.edu>)
Plate: acf01-10m3 row: d column: 10
Seq primer: M13F.

FEATURES
source
1..891
Location/Qualifiers
/organism="Asparagus officinalis"
/mol_type="mRNA"
/db_xref="taxon:4686"
/clone="acf01-10m3-d10"
/tissue_type="male inflorescences"
/lab_host="SOLR"
/clone_id="Acf01"
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
Site 2: XhoI; This is a directionally cloned,
non-normalized library. This library has been generated by

the Floral Genome Project (FGP). The Floral Genome Project
is funded by NSF's Plant Genome Research Program
(DBI-0115684). More information about the project can be
obtained at <http://fgp.bio.psu.edu>

ORIGIN
Alignment Scores:
Pred. No.: 2,34e-156 Length: 891
Score: 1374.00 Matches: 268
Percent Similarity: 97.58% Conservative: 14
Best Local Similarity: 92.73% Mismatches: 7
Query Match: 52.22% Indels: 1
Dbs: 7 Gaps: 0

US-10-718-952-12 (1-510) x CV290142 (1-891)

QY 222 ValAspIyValValIleuTrThrAlaAsnThrGluArgIySeriAsnLeuVal 241
Db 4 GTAGACAAAGTAGTGGTCTTTGGACTGCCAATACGAGAGCTATAGCAATGATTC 63
QY 242 GlyIyAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAla 261
Db 64 GGTCTCATATGACACATGAGAGACCTTGCTGCGCTGTGGACAAAGACAGACAGATA 123
QY 262 SerProSeriThrIyTrAlaIleAlaCysValMetGluAsnValProPheIleAsnGly 281
Db 124 TCCCATCAACCCCTATGATGATTCCTGCTGCTATGGAGAACCTTCCATCATCAACGGA 183
QY 282 SerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrIy 301
Db 184 AGCCCAAGAACACCTTCGCTGCTGTTGATTAATGGCTTATTCAAAGAAAGATCTTA 243
QY 302 IleGlyIyAspAspPheIySerylGlyInThrIyMetIySerValIleuValAspPhe 321
Db 244 ATGGAGAGATGACTTTAAGATGACAGACCAAGATGAAGTCGTGTGATTC 303
QY 322 LeuValGlyAlaGlyIleIyProThrSeriIleValSeriYTrAsnHisLeuGlyAsn 341
Db 304 CTGTGGGGGTGGAATTCMAACCACTCATATGTAGCTTCAACCATCTGGGAACAT 363
QY 342 AspGlyMetAsnLeuSeriAlaProGlnThrPheArgSeriIyGluIleSeriIySerAsn 361
Db 364 GATGCGATGATCTCTCGCACCAACCTTCGATCAAGAGAGATTCGAAGACAT 423
QY 362 ValValAspAspMetValAsnSeriAsnAlaIleuIyTrGluProGlyGluHisProAsp 381
Db 424 GTGGTCATGACATAGTTCAAAGCAATGCGATCTGTATGAGCTGGGAGACACCTGAT 483
QY 382 HisValValAlaIleIyTrValProIyTrValGlyAspSeriIyAspAlaMetAspGlu 401
Db 484 CATGCAATGTATCAAGATATGTCCTGATGTGTGTATACCAAGAGGCTATGATGAG 543
QY 402 TyrThrSeriGluIlePheMetGlyIyIyAsnThrIleValIleuHisAsnThrCysGlu 421
Db 544 TACACATTTGAGATATTCATGAGAGTCCAGCCCATTTGTCTGCACAAACCTGTGAG 603
QY 422 AspSeriLeuAlaAlaProIleIleIyAspLeuValIleuAlaGluLeuSeriThr 441
Db 604 GATTCGCTCTTGGCTGGCGCATCATACTGACTTGGTTATTTGGCTAGACCTTAGCAC 663
QY 442 ArgIleGlnPheIyAlaGluAsnGluGlyIyAspPheHisSeriPheHisProValAlaThr 461
Db 664 AGGATTCAGCTCAAGGCTGAGAGAGAGGCAAGTCCACTCTTCCTCCATCGTGGCTTACC 723
QY 462 IleIySeriTrIyLeuThrIyAlaProIyValProIyTrProIyValIleAsnAla 481
Db 724 ATCTTGATTAACCTCACCAAGGCTCCCTTGATTCACAGGACACCAAGTGTGAATGCA 783
QY 482 LeuSeriIyGlnAlaMetLeuGluIyAsnIleMetArgAlaCysValGlyLeuAlaPro 501
Db 784 CTCTCAAGCAAGAGGAGATCTGAGAAACATCTCAAGGCTTGCAATTTGGCTTGCTCT 843
QY 502 GluAsnAsnMetIleuGluTrIyAsp 510

DB 844 GAGAACAC-ATGATCCTGAGAGTACAA 869

Search completed: June 7, 2005, 22:49:35
Job time : 3751.5 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:22:56 ; Search time 219 Seconds
(without alignments)
3810.507 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631
Sequence: 1 MFIEFKVSPVVKYTTETI.....NIMRACVGLAPENNMLEYK 510

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10718952/rnatat.06062005.173401.12868/app_query.fasta_1.1294
-DB=Issued_Patents_NA -QFMT=fasta -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-IOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=humand40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORMexit -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952@cgn1_1.141@rnatat.06062005.173401.12868 -NCPU=6 -ICPU=3
-NO MMAP -LANG=EDUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONCLIG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfillseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2346	89.2	1931	3	US-09-118-442-10 Sequence 10, Appl
2	2346	89.2	1931	3	US-09-677-064-10 Sequence 10, Appl
3	2337	88.8	1931	4	US-09-727-628-10 Sequence 1, Appl
4	1823	69.3	3546	3	US-09-118-442-15 Sequence 15, Appl
5	1823	69.3	3546	3	US-09-677-064-15 Sequence 15, Appl
6	1822	69.3	3546	3	US-09-118-442-14 Sequence 14, Appl
7	1822	69.3	3546	3	US-09-677-064-14 Sequence 14, Appl
8	1389.5	52.8	1602	4	US-09-734-2378-72 Sequence 72, Appl
9	1389.5	52.8	1602	4	US-09-734-2378-72 Sequence 72, Appl
10	1384.5	52.6	1578	4	US-09-248-796A-6542 Sequence 131, Ap
11	1229.5	46.7	77626	4	US-09-949-016-12608 Sequence 12608, A
12	1103.5	41.9	1231	3	US-09-397-787-34 Sequence 34, Appl

c	13	529	20.1	34316	4	US-09-902-540-1257	Sequence 1257, Ap
c	14	528	20.1	1335	4	US-09-902-540-5256	Sequence 5256, Ap
c	15	160	6.1	294	4	US-09-313-294A-4684	Sequence 4684, Ap
c	16	142	5.4	42325	4	US-08-311-731A-131	Sequence 131, Appl
c	17	128.5	4.9	4403765	3	US-09-103-840A-2	Sequence 2, Appl
c	18	127.5	4.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl
c	19	119.5	4.5	1359	4	US-09-248-796A-6542	Sequence 6542, Ap
c	20	118.5	4.5	8930	4	US-09-077-098A-1	Sequence 1, Appl
c	21	118	4.5	7571	3	US-08-961-527-2	Sequence 2, Appl
c	22	117.5	4.5	274	3	US-09-118-442-21	Sequence 21, Appl
c	23	117.5	4.5	274	3	US-09-677-064-21	Sequence 21, Appl
c	24	116	4.4	2190	4	US-09-107-433-2104	Sequence 2104, Ap
c	25	116	4.4	2196	4	US-09-583-110-2059	Sequence 2059, Ap
c	26	115	4.4	2529	4	US-09-489-039A-4399	Sequence 4399, Ap
c	27	114	4.3	2193	1	US-09-014-897-3	Sequence 1, Appl
c	28	114	4.3	2193	1	US-09-014-897-3	Sequence 3, Appl
c	29	114	4.3	2193	1	US-08-731-716-3	Sequence 1, Appl
c	30	114	4.3	2193	1	US-08-731-716-3	Sequence 3, Appl
c	31	113.5	4.3	2337	4	US-09-489-039A-1284	Sequence 1284, Ap
c	32	112	4.3	3421	4	US-08-956-171E-337	Sequence 337, Ap
c	33	112	4.3	3421	4	US-08-781-966A-337	Sequence 337, Ap
c	34	110.5	4.2	2244	4	US-09-583-110-405	Sequence 405, Ap
c	35	110.5	4.2	2256	4	US-09-107-433-772	Sequence 772, Ap
c	36	110.5	4.2	5892	4	US-09-583-110-2582	Sequence 2582, Ap
c	37	110.5	4.2	5904	4	US-09-107-433-2280	Sequence 2280, Ap
c	38	110.5	4.2	640681	4	US-09-790-988-1	Sequence 1, Appl
c	39	109.5	4.2	1230230	4	US-09-438-185A-1	Sequence 1, Appl
c	40	109	4.1	8287	4	US-09-023-655-1441	Sequence 1441, Ap
c	41	108.5	4.1	3729	4	US-09-107-532A-1587	Sequence 1587, Ap
c	42	107	4.1	1052	3	US-08-961-527A-244	Sequence 244, Ap
c	43	106.5	4.0	2563	4	US-09-016-434-1076	Sequence 1076, Ap
c	44	106.5	4.0	2563	4	US-09-023-655-894	Sequence 894, Ap
c	45	106.5	4.0	2601	3	US-08-569-749-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10
Alignment Scores:
Pred. No.: 4,266-285
Score: 2346.00
Percent Similarity: 93.33%
Best Local Similarity: 87.84%
Query Match: 89.17%
DB: 3
Length: 1931
Matches: 448
Conservative: 28
Mismatches: 34
Indels: 0
Gaps: 0

US-10-718-952-12 (1-510) x US-09-118-442-10 (1-1931)

QY 1 MetPheIIeGluAspPheIValIglUSeRProAsnValIleYThrGluThrGluIle 20
Db 99 AAGTTCATGAGAGCTTCGCGTCGAGAGCCCCCGGAGCGGCCGATGAGATC 158
QY 21 GlnSerValIYrAsnIYrGluThrThrGluLeuValIleGluAsnArgAsnIYrThyTyr 40
Db 159 GAGTGGAGTACCGGATGACAGACAGAGAGCTGTATCAGAGAGGCGAAGAGACGGCGCTCA 218
QY 41 GlnTrpIleValIYrAsnIYrProIYrSerValIAsnIYrGlnPheIleThrAsnThiIValPro 60
Db 219 CCTGGGTGTCGCCGCCCAAGTCCGTCAGATCAACTCCGAGACCAAGACCCCGCTCCC 278
QY 61 LysLeuGluValMetLeuValIglYTrpGlyIAsnAsnGlySerThrLeuThrGlyGly 80
Db 279 AAGCTCGGGGTAGTGTGTGGGGTGGGAGGCAACAGGGGTCCACGCTGACGGCTGGG 338
QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 339 GTTCATTTGCCAAGAGGAGGAGATCTCATGGGCAACCAAGACAGGTGACGACCAAC 398
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValIglYSerPheGlnIglUguIu 120
Db 399 TACTACGGCTCCCTCCACCGCTCCACATCAGAGTGGGAGCTCAACAGGGAGAG 458
QY 121 IleYrAlaProPheIYrSerLeuLeuPProMetValAsnProAspIleValIlePheGly 140
Db 459 ACTTAAGCCCGCTTCAAGAGCCCTCTTCCATAGTAAGAACCCAGACCACTTGTGTGGA 518
QY 141 GlnYrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIAsnValIlePheAsp 160
Db 519 GCGTGGAGCATTTAGCAACATGAACCTGGCGCATCATATCAAGGCGCAAGGTGCTGAT 578
QY 161 IleAspLeuGlnLysGlnLeuArgProIYrMetGluSerMetValIProLeuProGlyIle 180
Db 579 ATTGACCTGACAGAGAGCTCAGGCGCTCATGAGATTCATGATGTCCTCCCGGTATC 638
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnIglUArgAlaAsnAsnValIleIleGly 200
Db 639 TATGATCCCGACTTCATCCGCGCTAACCGAGGCTCGCGCCACAGTGTCTATCAAGGGC 698
QY 201 ThrLysGlnIglUValIleGlnIleIleLysAspIleLysAlaPheIleGluAlaThr 220
Db 699 ACCAAGAAAGAACAGGTGAGACGATCATCAAGATATACGGAGATTTAAGAGAAAGAAC 758
QY 221 LysValIAspLysValIValIleuTrpThrAlaAsnThrGluArgIYrSerAsnLeuVal 240
Db 759 AAAGTGCACAAAGATAGTGTGTGTGTGACTGCAGAAACACTGAAGATATGCAATGTGTC 818
QY 241 ValIglYLeuAsnAspThrMetGluAsnLeuLeuAlaValIAspArgAsnIglUArgIu 260
Db 819 GCTGGCTCAAGACACGATGAGATCTACGTGGACTGTGAGACAAAGAACAGCGGAG 878
QY 261 IleSerProSerThrLeuYrAlaIleAlaCysValMetGluAsnValIProPheIleAsn 280
Db 879 GATACCACTCAACACTATATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 938
QY 281 GlySerProGlnAsnThrPheValIProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 939 GGGAGGCCCCAGAACACTTTGTGCTGGCTGATTGATCTTGCTATATAAAAAACAACCTGC 998
QY 301 LeuIleGlyGlyAspAspPheIYrSerGlyGlnThrLysMetLysSerValIleuValAsp 320
Db 999 TTGATTTGGTGTGACACACTTCAAGAGTGCAGACAGACCAAAATGCTCTTGTGTGCTAT 1058
QY 321 PheLeuValIglYAlaGlyIleLysProThrSerIleValSerTrpAsnIleuGlyAsn 340
Db 1059 TTCTCTTTGTTGGCTGGAATTAAGCCCACTCATGTGATCAACCAACTTGGGAAAC 1118
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360

Db 1119 AACATGGCATGAACCTGTCTGCCCTTCAACATTCAGCTCCAAAGAGATCTCCAAAGAC 1178
QY 361 AsnValIleAspAspMetValIAsnSerAsnAlaIleLeuYrGluProGlyGluIleAspPro 380
Db 1179 AAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1238
QY 381 AspIleValIleValIleLysYrValIProYrValIglYAspSerLysArgAlaMetAsp 400
Db 1239 GATCATGTGTTGTATCATGAATGTGTCGTATGCGGAGAGACAGCAAGAGGGCTATGAC 1298
QY 401 GlnYrThrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuIleAsnThrYs 420
Db 1299 GAGTACACCTCAGAGATCTTATGAGCGGCAAGAACACATCTGTGCTGCACAAACCTGT 1358
QY 421 GluAspSerLeuLeuAlaAlaProIleIleuAspLeuValIleuValIglUleuSer 440
Db 1359 GAGAGCTGCTCTCGCCGACCTATCTCTTGTATCTGTGTGTGTGTGTGTGTGTGTGT 1418
QY 441 ThrArgIleGlnPheLysAlaGluAsnGluIYrLysPheIleSerPheIleProValAla 460
Db 1419 ACCAGATCCAGCTGAAGAGCTGAGGAGAGAGCAAAATTCATCTCTTCCACCCGGTGGC 1478
QY 461 ThrIleLeuSerYrLeuThrLysAlaProLeuValIProProGlyThrProValIleAsn 480
Db 1479 ACCATCTTGAAGTACTTCCACCAAGGACACCTGTGTCCCTGTGGACACCGGTGTGAC 1538
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValIglYleuAla 500
Db 1539 GCTCTGGCCAAAGCAGAGGCGCATGTGTGAGAAACATCATGAGAGGCTGTGTGTGTGTGT 1598
QY 501 ProGluAsnAsnMetIleLeuGluYrLys 510
Db 1599 CCAAGAAACAAATGATCTTGAAGTACAG 1628

RESULT 2
US-09-677-064-10
; Sequence 10, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytochemical Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-10

Alignment Scores:
Pred. No.: 4,26e-285 Length: 1931
Score: 2346.00 Matches: 448
Percent Similarity: 93.33% Conservative: 28
Best Local Similarity: 87.84% Mismatches: 34
Query Match: 89.17% Indels: 0
DB: 3 Gaps: 0

US-10-718-952-12 (1-510) x US-09-677-064-10 (1-1931)

QY	1	MetPheIleGIuaenPheIyValGIuseProaenValySYrThrGluThGluIle	20
Db	99	ATGTCATCGAGAGGCTCCGGTCGAGAGCCCCCAAGCGGTACGGCCCATGAGATC	158
QY	21	GIuseValTYraenTYrGIuThrThrGIuLeuValIHISGuaenArGIYThrTYr	40
Db	159	GAGTCGAGATACCGGTACGACACGAGCTGTACACGAGGGCAAGACGCGCCCTCA	218
QY	41	GIuThrPIleValIyEProIySereValaenTYrGIuPheIyThrIaenThrHISValPro	60
Db	219	CGCTGGGCTCGTCCGCCCAAGCTCGTCAAGTACAATCTCCGGAACGACCGCGTCCCC	278
QY	61	LYseIuGIyValMetLeuValGIYThrPIyGIYaaenAnGIYserThrIeuthGIyGIY	80
Db	279	AAGCTCGGGGAGAGCTTGCGGGGAGGACCAACGGGTCCACGCTACCGCTGGG	338
QY	81	ValIleIaaenArGIuAspIleSerThrPIaThrIyAspIyIleGIuIuIaasn	100
Db	339	GTCATTCGCCACAGGAGGGAGTCTCAIAGGCGACCAAGACAGATGTCAACGCCAC	398
QY	101	TYrPheGIYserIeuthThrIuIaIaSerAlaIleaYValGIYserPheGIuGIYGIuIu	120
Db	399	TACTACGGCTCCCTCACCGCTCCACATCAGATGCGGAGCTTACACGGGGAGAG	458
QY	121	IIeTYrAlaProPheIySerIeuleuProMetValaenProAspIleValPheGIY	140
Db	459	ATCTATGGCGCGTTCMAAGGCTCTCTCCATAGGAACCAAGACATGTGTTCGGA	518
QY	141	GIYThrAspIleSerAsnMetAsnIeulIaAspAlaMetAlaYrAlaIyValPheaSP	160
Db	519	GGCTGGGACATTAGCAATGAACTTGGCCGACTCATTGACACAGGCGCAAGGTGCTGAT	578
QY	161	IIeAspIeGIuIyGIuIeuaYrProTYrMetGIuseMetValProIeuproGIYIle	180
Db	579	ATTGACCTCGAGAAGCACTCAGGCGCTTACATGAGCTCATGATGCCACTTCCGGTATC	638
QY	181	TYrAspProAspPheIleAlaIaasnGIuGIuIaYrAlaAsnAsnValIleIyGIY	200
Db	639	TATGATCCGGACTTCATCGCGGCTTACACGGGCTCTCCGCGCAACAGTGTATCAAGGGC	698
QY	201	ThrIySGInGIuGIuIaGIuIleIleIyAspIleIyAlaPheIyGIuIaIaThr	220
Db	699	ACCAAGAAAGAACAGGTGAGGAGATCAAGGATATCAGGAGATTAAAGAGAAAGAC	758
QY	221	LYsValAspIyValValIleuThrThrAlaenThrGIuIaYrSerAsnIeulVal	240
Db	759	AAAGTGGACAAAGATGTGTGTGTGTGGACTGCCAAACCTGAAGATTAAGCAATGTGCG	818
QY	241	ValGIyIeuaenAspThrMetGIuAsnIeuleuAlaIaValAspArGIuIaGIu	260
Db	819	GCTGGTCTCAACGACACATGAGAGATTACTGCGCATGTGGACAAAGAACGAGCGGAG	878
QY	261	IIleSerProAspThrIeuthAlaIleAlaCySValMetGIuaenValProPheIeIasn	280
Db	879	GTATCACCATCAACATATATGCCATTGCTGTGCATGAGAGGGGTCCGTTCACTCAAT	938
QY	281	GIYserProGIuAsnThrPheValProGIYIleuIleAspIeulAlaIleaYrAsnThr	300
Db	939	GGAGGCCCCAGAACACTTGTGTGCTGGCTGGCTGATGATCTTGTCTATTAATAAACATCGC	998
QY	301	LeuIleGIYIyAspAspPheIySereGIYGIuThrIyMetIySereValIeulValAsp	320
Db	999	TTGATGTGTGTGACGACTTCAAGAGTGGACCAAGACATGAATCTGTCTTGGTGGAT	1058
QY	321	PheIeulValGIYAlaGIYIleYrProThrSerIleValaSerTYrAsnHISleuGIYasn	340
Db	1059	TTCCTGTGTGGTGGGAATTAAGCCCACTCAACGTGAGCTTACACACTTGGGAAC	1118
QY	341	AsnAspGIYMetAsnIeulSerAlaProGIuThrPheArGIYserIyGIuIleSerIySere	360
Db	1119	AAACATGGCAATGACCTGTCTGCCCTCAACATTCACAGTCCAGAGATGTCCAAAGGC	1178
QY	361	AsnValValAspAspMetValaenSerAsnAlaIleuTYrGIuProGIYGIuIaIaPro	380

Db	1179	AACTGTGGATGATGATGCTCGACGAATCCATCTCTATGAGCCCGGAGCATCCC	1238
Qy	381	Asph1sValValVal111eLyTyTyValProTyValG1yAspSerLySargAlAmelAsp	400
Db	1239	GATCATGTGCTGTGCATCATAGTATGTGCCGAGCGTGGAGACAGCAAGGGCTATGAGC	1298
Qy	401	GIuTyThrsSerGIu11ePheMetGIyLyAsnThrl1eValLeuH1sAsnThrCys	420
Db	1299	GAGTACACCTTCAGAGATCTTCATGTGGGGGAGAACCAATCGTGTGCACAAACACTGT	1358
Qy	421	GIuAspSerLeuLeuAla1aPro11e1eLeuAspLeuVal1eLeuAlaGIuLeuSer	440
Db	1359	GAGAGCTGCTCTCTCGCCGACCTTATCTTGTATCTGTGTGCTCTTGCGCTGAGCTAGC	1418
Qy	441	Th1rArg11eG1nPhelysAlaGIuAsnGIuLyLyAspPheH1sSerPheH1sProValAla	460
Db	1419	ACCAAGATCCAGCTGAAAGCTGAGGAGGAGAGAACAAATTCACATCTTCCACCCGGTGCC	1478
Qy	461	Thrl1eLeuSerTyTyLeuThrTyTySAlProLeuValProProGI1yThrProValValAsn	480
Db	1479	ACCATCTTGAGTACTTCACCAAGGCAACCCCTGGTTCCTCCGACACCGGTGTGTAAC	1538
Qy	481	AlaLeuSerLySGIuSargAlAmelLeuGIuAsn11eMetArgAlaCysValGIyLeuAla	500
Db	1539	GCTGTGGCAAGCAGAGGGCGATGCTGAGAACATCAATGAGGCGCTGCGTGGGCTGGCC	1598
Qy	501	ProGIuAsnAsnMet11eLeuGIuTyTyLyS	510
Db	1599	CCAGAGAACCAATGATCTTGAGCTACAG	1628
RESULT 3			
US-09-727-628-1			
; Sequence 1, Application US/09727628			
; Patent No. 6791013			
; GENERAL INFORMATION:			
; APPLICANT: Armstrong, Katherine			
; APPLICANT: Hey, Timothy D			
; APPLICANT: Folkerts, Otco			
; APPLICANT: Smith, Kelley A			
; APPLICANT: Hopkins, Nicole L			
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER			
; FILE REFERENCE: 50597			
; CURRENT APPLICATION NUMBER: US/09/727,628			
; CURRENT FILING DATE: 2000-12-01			
; PRIOR APPLICATION NUMBER: US 60/168,612			
; PRIOR FILING DATE: 1999-12-02			
; NUMBER OF SEQ ID NOS: 3			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 1959			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (137)..(1699)			
US-09-727-628-1			
Alignment Scores:			
Pred. No.: 5.97e-284			
Score: 2337.00			
Percent Similarity: 93.33%			
Best Local Similarity: 87.84%			
Query Match: 88.83%			
DB: 4			
Gaps: 0			
US-10-718-952-12 (1-510) x US-09-727-628-1 (1-1959)			
Qy	1	MePhe11eG1uAsnPhelysValGIuSerProAsnValTyTyThrGIuThrGIu11e	20
Db	137	ATGTTCATTCAGAGACTTCCGCGTGAAGACCCCAACGATGCGGThrAGCCGACAGCATC	196
Qy	21	G1nserValTyTyAsnTyTyGIuThrThrGIuLeuValH1sGIuAsnArgAsnGIyThrTy	40


```
Db 197 GAGTGGAGTACCGGTACGACGACGAGGCTGGTGCACAGAGGCGCAAGGACGGCGCTCC 256
Qy 41 GlnTPrIleValPProLySerValAsnTrpGlnPheLySThrAsnThiSValPro 60
Db 257 CGCTGGGTCGCTCCGCCCAAGTCGTCAGTCAACTTCCGGACCAACCGCGGTCC 316
Qy 61 LysLeuGlyValMetLeuValGlyTrpGlyAsnAsnGlySerThrLeuThrgly 80
Db 317 AAGCTCGGGGTCTATCTTGTGGGTGGGAGGCAACAGGCTCAACCTGACGGCTGG 376
Qy 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 377 GTCATTCGCCAACAGGAGGAGGATCTCATGGGCGCAACAGAGCAAGAGTGCAGCAACCAAC 436
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlnGlu 120
Db 437 TACTACGGCTCCCTCACCAGGCTTCACCATCAAGATGAGCGAGTCAACCGGGAGAG 496
Qy 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 497 ATATATGGCCGCTTCAAGAGCTCTTACCTCATGTGTAACCAACGACCTTGTGTTTGA 556
Qy 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 557 GGCTGGACATCAGACGATGAACCTGGCAGATGCATGACAGGAGCCCAAGGCTGGAC 616
Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProLysIle 180
Db 617 ATTGACCTGCAGAAACAGCTGAGGCTCATGAGTGCATGAGTGCCTCCGCTGTC 676
Qy 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnValIleLysGly 200
Db 677 TATGATCCGAGCTTATCGCCCTAACCAAGGCTCTCGGCCAACATCTCATAGGCG 736
Qy 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db 737 ACCAAGAAAGAACAGGTGAGGACATCATCAAGATATCAGGAGATTAAAGAGAAAGAAC 796
Qy 221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgTyrSerMetLeuVal 240
Db 797 AAAGTGACACAGGTGTGTGTGTGTGTGACTGCACAACTGAAAGGTACAGCATATGAT 856
Qy 241 ValGlyLeuAsnAspThrMetGlnLeuLeuAlaValAspArgAsnGlnAlaGlu 260
Db 857 GGTGGTCTCAACGACATGAGATCTGTGGCATCTGTGGAACAAGACAGGCGGAG 916
Qy 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 917 ATCTCGCCATCAACACTATATGCATGTGCTGTGTCAAGGAGGCTGCGTTCATCAAT 976
Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 977 GGGAGCCCCCAGAACCTTTGTGCTGGCTGATGATCTTGATCTTCAAGAACCAACTGC 1036
Qy 301 LeuIleGlyGlyAspAspPheLysSerGlyThrLysMetLysSerValLeuValAsp 320
Db 1037 CTGATCGGAGGAGGACACTTCAAGAGTGGGACAGACCAAGATGAATCGTCTGTTGAT 1096
Qy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleGlnGlyAsn 340
Db 1097 TTTCTTTGTTGGTGTGAGTAATAAGCCACCTCGATTGTAGCTCAACACCACTTGGGAAAC 1156
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db 1157 AACGACGGGATGAACCTGTCTCCCTCAAACTTCAGTCCAGAGAGTCTCCCAAGAC 1216
Qy 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnHisPro 380
Db 1217 AACGGTGGTGAATGACATGCTTCACAAATGCAATTCCTTATGGGCCCGGACGATCCC 1276
Qy 381 AspHisValAlaValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
```

```
Db 1277 GATCATGTTGTTCATCAAGTATGTGCCGTATGTGGAGACAGTAAAGGCTATGAC 1336
Qy 401 GlnTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
Db 1337 GAGTACACATCAGAGATCTTATGGGCGGCAAGACACATCTGTGTCAACAACCTGC 1396
Qy 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1397 GAGAGCTGCTCTTCGCCGACCGATCATCTCGATCTGGGTCTGTGGCTGAGCTCAGC 1456
Qy 441 ThrArgIleGlnPheLysAlaGluAsnGluLysPheHisSerPheHisProValAla 460
Db 1457 ACCAGGATCCAGTTAAACCTGAGGGAACGACAAAGTTCCTCTTCCACCGGTTGCC 1516
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValAlaAsn 480
Db 1517 ACCATCTTACTACTTACCTACCAAGGACCACTGTGTCCACCGGACACCGGTGTGAC 1576
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1577 GCTCTTGCAAGCAGAGGCGATGCTGGAGAACTCATGAGGGCTTGGCTGGCTGGCC 1636
Qy 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1637 CCAGAGAACCAACATGATCTGAGTACAG 1666
```

RESULT 4

```
US-09-118-442-15
/ Sequence 15, Application US/09118442B
/ Patent No. 6197561
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Calt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ APPLICANT: Wang, Xun
/ APPLICANT: Bowen, Benjamin A.
/ TITLE OF INVENTION: Genes Controlling Phycate Metabolism in
/ TITLE OF INVENTION: Plants and Uses Thereof
/ FILE REFERENCE: 0706
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ CURRENT FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-118-442-15
```

Alignment Scores:

Pred. No.:	1,16e-218	Length:	3546
Score:	1823.00	Matches:	446
Percent Similarity:	40.86%	Conservative:	28
Best Local Similarity:	38.45%	Mismatches:	36
Query Match:	69.29%	Indels:	652
DB:	3	Gaps:	9

US-10-718-952-12 (1-510) x US-09-118-442-15 (1-3546)

```
Qy 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyrThrGluThrGluIle 20
Db 60 ATGTTCAATCGAAGAGCTTCGCGTCAAGAGCCGCCAGCGTACGCGCCACGAGATC 119
Qy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 120 GAGTGGAGTACCGGTACGACGACGAGGCTGTGTACAGAGGAGCAAGAGACGCGCTCA 179
```

QY 41 GlnTPIleValIysProIysSerValIasnYrIghPheIysThrAsnThrH1sValPro 60
Db 180 CCGTGGTGTGCGCCGCAAGTCCGTCAAGTACAACTTCGACCAAGACCGCCGTCCCC 239
QY 61 LysLeuGly----- 63
Db 240 AAGCTCGGGTANGTAGAGATGACGGGGCCCTAGCCTCACTCTGTGAAACCTCCCTCTC 299
QY 63 ----- 63
Db 300 CCGTCTCAGTCAAAATCCTCCGTGAGATCAACTGGTCGCGCTCCCTCTAAATCTTAA 359
QY 63 ----- 63
Db 360 TGAATAATCTTACTGCTTGGCTTGAAAGACGACCGTGTATTTGTGACAGCTACGCAAC 419
QY 63 ----- 63
Db 420 ACTTGCCCATCGGGATGGGTCAAAATCAGCTGATTTGAAATTCGATGATGGTCCCTT 479
QY 63 ----- 63
Db 480 TTCCATATTTTGCATCATCTTCGCTACTGTCGAATGATTAGAGAAACGTCCTTCTCTC 539
QY 63 ----- 63
Db 540 TGAATCTTGTCTTAAAGCTTTTGTCTGTGACGTGAGCTGTGTATTCATTTGTTCATGTA 599
QY 63 ----- 63
Db 600 AGATCAAAATTCACGACGAGGACGATGACGACGACAGACAGAACTATTACGTACGAAATTGA 659
QY 63 ----- 63
Db 660 TACTAGATTACTGGCAAGTGTGATACGCGCGCAATCTGCATCTGACCCCTTTGTT 719
QY 63 ----- 63
Db 720 TAATTCCTGTCTTATGCAATGTTGCTTACGTGACGTGTTGTGTATGATGTCAGGC 779
QY 63 ----- 63
Db 780 TGTCAAGCCGCTTGTCTCTGTCTGACGAGATGATCCAACTTTCTGTTCGTGGTGCAGG 839
QY 64 ValMetLeuValIglYTrpGlyIysAsnAsnGlySerThrLeuThrGlyValIleAla 83
Db 840 GTGATGTCTTGTGGGTGGGAGGACAAACGCGGTCCAGCTGACGCGCTGGGTCAATTGCC 899
QY 84 AsnArgGlu----- 86
Db 900 AGCAGGAGTGAAGTAGTACTTATTTGTCTTATTTGCTTTCGTTGTTTTCAGTTATTA 959
QY 87 ----- 92
Db 960 ATGGCCCTGACAGAACTGAATTTTGTGTGTGGCTGTTTCAGGGGATCTCAGGCCGAC 1019
QY 93 LysAspLysIleGlnGlnIleAsnYrPheGlySerLeuThrGlnIleAsnIleAlaIleArg 112
Db 1020 AAGGACAAAGGTGACAGAACCACTACACGGCTC-CTCACCCAGGCTCCACACATCAGA 1078
QY 113 ValIglySerPheGlnGlyIleGlnIleIleYrAlaProPheIysSerLeuProMet--- 131
Db 1079 GTGCGACGCTACACGCGGAGAGATCTATGCGCGTTCAGAGGCTCCTTCCCATGTGTA 1138
QY 131 ----- 131
Db 1139 ATCTATTATAGCTTGACTTAATCTCTTTTACTGAAACCAAACTATCATTAACAAG 1198
QY 131 ----- 131
Db 1199 CATATTCCGTAAAGGTCTAGTTGATGTATTAAGTGAACCTGTCTTTCAGGCCAGTGTTC 1258
QY 131 ----- 131

Db 1259 TCAAGTAAACGGAATGTTAATCATTTGGGTGAAAAAACAAAGTTCTAATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAATGTTAACTTAGCATATATGAAAAAGGGAAGCATTTGTAGAAAAAGGTGTGAACAATC 1378
QY 131 ----- 131
Db 1379 GACTCGGTCTCATGTGTGTGATCTTACTTGCAAGTCAAAAAGTTCTGTGTAGCTCAA 1438
QY 131 ----- 131
Db 1439 AGGTTCCAGCATCTTTGATTTACACTGCTGACGATTGACGATGGTGTAACTGGTTGCA 1498
QY 131 ----- 131
Db 1499 GATTGACAGACTCGGTGTTGTATCTTCTTTCATGACCAAGTGTGMACTGTTTCA 1558
QY 132 -ValAsnProPheAspIleValIlePheGlyIleTrpAspIleSerAsnMetAsnLeuAlaAs 151
Db 1559 GGTGAACCCAGACGACATTTGTGTTCGAGGCTGGACATTAGCAACATGAACTGGCCGA 1618
QY 151 PALawerAlaArgAlaIysValIlePheAspIleAspLeuGlnIleuArgProTyMe 171
Db 1619 CTCCATGACAGGCGCAAGGTCTGATGATTTGACTGCAAGACAGCTCAAGGCCCTACAT 1678
QY 171 fGluSerMetValIProIeu-ProGlyIleYrAspProAspPheIleAlaIleAsnGln 191
Db 1679 GAGTGTCAATGGTGCACATTTCCCGGTATCTATGATCCGGACTTCACCGGCTTAAACAG 1738
QY 191 IuGluArgAlaAsnAsnValIleIysGlyThrIysGlnIleGlnIleAlaIleIle 211
Db 1739 GCTTCGCGCCAAAGTGTCTATCAAGGCGCAACAAAGAAACAGGTGACGAGATCA 1798
QY 211 YAspIle----- 213
Db 1799 AGGATATCAGGTATATGATATGATGCTTAAACGTCCTTGGTCTTAAGTGCACCCAGTG 1858
QY 213 ----- 213
Db 1859 CAACCTAAAGAAATTAATTAATAATTTGTTGTAATATATACATATACAGAACAT 1918
QY 213 ----- 213
Db 1919 ATTTGTTAACCGGTGACTTAGAAGTCTGATGATGTGGACAAATTTGACATTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGCTCATCTTGACATGAGACTCCACAAAGAACTTAAACTTAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTGCTAGCTAATGATGATTTGTGTGACATGGTTTGAAGATCTAGATTAAAGT 2098
QY 213 ----- 213
Db 2099 GATTGAAGAAATATGTTCACTAGACATAAGTAATCCATTACAGAAAGCTATGGCTTATGTA 2158
QY 214 -----LysAlaIlePheIysGlnIleAlaThrIysValIleAspLysValIleValIleuTP 230
Db 2159 GCTTAATGAACAGGAGTTTAAAGAGAAACAAAGTGAACAAAGTATGTGTGTGGA 2218
QY 230 hrAlaAsnThrGluArgYrSerAsnLeuValIleGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACATGAAAGGTATGCAATGTGTGCGCTGCTCAGACACGATGAGATC 2278
QY 250 euLeuAlaIleValAspArgAsnGlnIleGlnIleSerProSerThrLeuYrAlaIle 270
Db 2279 TACTGCAATCTGTGGAACAAGACGAGGAGATATCACATCAACATATATGTCATTTG 2338
QY 270 IAcYsValMetGluAsnValIleProPheIleAsnGlySerProGluAsnThrPheValIPro- 289

```
Db 2339 CCTGTGTCAATGAGGGGGGCGCCGTTTCATCATGAGGAGCCCGACAGACACCTTTGTGCTG 2398
QY 289 ----- 289
Db 2399 GTGCGGTGTGTGTGTGTGTGAAAAAGCTTCANGGTGTTCATTTCTGTTCAAAGTTTC 2458
QY 290 -----GlyLeu1le 292
Db 2459 ATGGTGTGTATTTCCGTTCCAGGCTTATTTACCTGTGTGATGTCGTGAGGCTGATTT 2518
QY 293 AAPPLeuAla1AlaArgAsnThrLeu1leGlyAspAspPheLysSerGlyGlnThr 312
Db 2519 GATCTTGTCTATTAATAAAACACTGCTGATTTGGTGGTGAAGACTTCAGAGTGAAGACACC 2578
QY 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGly1leLys----- 328
Db 2579 AAGATGAATACTGTCTGTGTGCAATTTCTTGTGTGGTGGGAAT-AAAGGTGGAAACCTA 2637
QY 328 ----- 328
Db 2638 GTATCTCTCTCTATTAAGATGAAGTGTGTTTGTGGCAATGACGTTATTGCAATTAATCTC 2697
QY 329 -----ProThrSer1leValSerTyrAsnHisLeuGlyAsn 340
Db 2698 TTCTATATTTCATTTTCATGACAGCCCACTCAATCGTAGCTTCAACACACTTGGGAAAC 2757
QY 341 AanaAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGln1leSerLysSer 360
Db 2758 AACGATGGGATGAACCTGCTGCGCCCTTCAACATTCAGGTCCAAAGAGATCTCCAAAGAC 2817
QY 361 AenValValAspAspMetValAsnSerAsnAla1leLeuTyrGlnProGlyGlnHisPro 380
Db 2818 AACGTGGGTGATGAATGCTGTCTGAGCAATGCATCTCTATGAGACCCGCGCAGATCC 2877
QY 381 AspHisValVal1leLys----- 387
Db 2878 GATCATGTCGTTGTATCAAGGCTCTGTTAGCTGATCTTCACTGCTTAAAGATGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGGACGATTTACATGAAACCTGTCACTCTTTTGTGCAATAGTGTGCGTAACTGT 2997
QY 393 GlyAspSerLysArgAlaMetAspGluTyrThrSerGln1lePheMetGlyGlyLysAsn 412
Db 2998 GGAGGACAGAGAGAGGCTATGACAGATCACTCAAGATCTTCATGGCGGCAAGAAC 3057
QY 413 Thr1leValLeuHisAsnThrCysGlnAspSerLeuLeuAla1Pro1le1leLeuAsp 432
Db 3058 ACCATCGTGTGCAACAACCTGTGAAGACTGCTCTCGCCGACCTATCATCTTGAT 3117
QY 433 LeuVal1leLeuVal1leGlyLeuSerThrArg1leGlnPheLysAlaGlyAsnGlu----- 450
Db 3118 CTGGTGTCTTGGCTGAGCTGACACAGATCCAGATCGAAAGCTGAAGGAGAGGTAAAGA 3177
QY 450 ----- 450
Db 3178 GCCCCCAAGATTAACCTGAAGACAGCTGACAGCTAGGTGATATAGCACTTTAATA 3237
QY 451 -----GlyLysPheHisSerPheHisProValAla1leThr1le 462
Db 3238 CTTTCTGGTGTCTCTTATGACAGACAAATTCACCTCTTCCACCGGTGGCCACATTC 3297
QY 463 LeuSerTyrLeuThrLysAlaProLeu----- 471
Db 3298 CTGACCTACCTCAACCAAGGCAACCTGTGAAGCTTTTCTCTGATCCCGGACATCACTG 3357
QY 472 -----Val 472
Db 3358 CACTGCTTTTGTCTTCAATCCAGCACTGATGCTCTCTCGAAACCTGAACAACAGGTT 3417
QY 473 ProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGlnAsn1le 492
Db 3418 CCCCTGGCAACACCGGTGTGTAAAGCTCTGTGCAAGACAGACCGCGATGCTGTGAAGAACATC 3477
```

```
QY 493 MetArgAlaCysVal1leGlyLeuAlaProGlnAsnAsnMet1leLeuGlyTyrLys 510
Db 3478 ATGAGGGCGCTCGTTGGGCTGGCCCGACAGAAACAATGATCTGAGTACAAG 3531

RESULT 5
US-09-677-064-15
; Sequence 15, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Galt, Susan J.
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-15

Alignment Scores:
Pred. No.: 1.16e-218 Length: 3546
Score: 1823.00 Matches: 446
Percent Similarity: 40.86% Conservative: 28
Best Local Similarity: 38.45% Mismatches: 36
Query Match: 69.29% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-12 (1-510) x US-09-677-064-15 (1-3546)
QY 1 MetPhe1leGlnAsnPheLysVal1leGlyLeuVal1leSerProAsnVal1leTyrThrGln1le 20
Db 60 ATGTTTCATCGAGACTTCCGGTGTGAGAGCCCCACGTCGCGTACGGGCCGACGAGATC 119
QY 21 GlnSerVal1leTyrAsnTyrGlnThrThrGlnLeuValHisGlnAsnArgAsnGlyThrTyr 40
Db 120 GAGTGTGAGGTACCGGTACGACACAGACGAGCTGTACACGAGGACGAGGAGGCGCTCA 179
QY 41 GlnThr1leVal1leLysProLysSerValAsnTyrGlnPheLysThrAsnThrHisValPro 60
Db 180 CGCTGGGTCTGCGCCCAAGTCCGTGAAGTCAACTTCGGACCAAGACCGCGTCCC 239
QY 61 LysLeuGly----- 63
Db 240 AAGCTCGGATATGTAACGATGACAGGCGCCCTAGCTCTCTGTGAACCCCTCTCTC 299
QY 63 ----- 63
Db 300 CCGTGTCACTCAATCTCCGTGAGATGAACGTGCGGTTCCCTCTTAATCTTAA 359
QY 63 ----- 63
Db 360 TGAATCTTACTGCTTTCCTGAAGACAAACGTCGTAATTGTGAACAGCTACGACAC 419
QY 63 ----- 63
Db 420 ACTTGCCCATCCGAGTGGCTCAATTCAGCTGATTTGAATTGATTCGATTCGATTC 479
QY 63 ----- 63
```

[illegible]

```
QY 329 -----ProThSerIleValSerTyraSnhIleuGlyAsn 340
Db 2698 TTCTATTATTTTCATTTTCATGCAGCCCACTCAATCGAGACTACCAACCTTGGAAGC 2757
QY 341 AsnAspGlyMetAsnLeuSerLysProGlnThrPheAspSerLysGlnIleSerLysSer 360
Db 2758 AACGATGGCAGTGAACCTGCTGCGCCCTTCAMACATTCAGGTCGAGAGATCTCCAAAGAC 2817
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyGluProGlyGluHisPro 380
Db 2818 AACGTGGTGGATGAGATGGTCTCGAGCAATGCCATCTTATGAGCCCGGCGAGCATCCC 2877
QY 381 AspHisValValValIleLys-----TyValProTyVal 392
Db 2878 GATCATGTCGTGTGATCATGAGGTCTGTAGCTGATCTTACCTGTTAAAGTTGACAT 2937
QY 388 -----TyValProTyVal 392
Db 2938 ATGCAAGCGAGATTACATTGAACTGTCACTCTTTGTCAGATGATGTGCCGTACCTG 2997
QY 393 GlyAspSerLysArgAlaMetAspGluTyThrSerGluIlePheMetGlyGlyLysAsn 412
Db 2998 GGAGACAGCAGAGAGGCTATGAGACAGTACACCTCAAGATCTTCAAGGCGGCAAGAC 3057
QY 413 ThrIleValIleuHisAsnThrCysGlnAspSerLysLeuAlaAlaProIleIleuAsp 432
Db 3058 ACCATCGTGTGCACACACACCTGTGAGGACTGCTCTGCGCGCACTATCATCTCTGAT 3117
QY 433 LeuValIleuLeuAlaGluLeuSerThrArgIleGlnPheLysValGluLysGlu----- 450
Db 3118 CTGGTGTCTTGTGGCTGAGCTCAGACCAAGATCCAGCTGAAGCTGAGGAGAGGTAAAG 3177
QY 450 ----- 450
Db 3178 GCCCCCCAAGTATTACCTGAAGACGCTCAGGCTAGGTGATGATGACATTTTAATA 3237
QY 451 -----GlyLysPheHisSerPheHisProValAlaThrIle 462
Db 3238 CTTTGTGGTGTCTCTTATATGAGGACAAATTCACCTCTTCCACCGGTGGCCACATC 3297
QY 463 LeuSerTyLeuThrLysAlaProLeu----- 471
Db 3298 CTGAGCTACCTCACCAAGGCAACCTGTGAAGCTTTCTCTGATCCCGGACATCACTG 3357
QY 472 -----Val 472
Db 3358 CACTGCGTTTCTTCAATCCAGCACTGATGCTCTCTCGAAACCTGAACAACAGTT 3417
QY 473 ProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGluLysIle 492
Db 3418 CCCCCTGGCAGACACCGGTGTGAACGCTCTGCGCAAGCAGACGCGATGTGAGAAACATC 3477
QY 493 MetArgAlaCysValGlyLeuAlaProGluAsnAsnMetIleLeuGluTyLys 510
Db 3478 ATGAGGGCGCTGCTGGCTGGCCCGAGAGAACATGATCTCTGAGTAAAG 3531

RESULT 6
US-09-118-442-14
; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Wang, Larry R.
; APPLICANT: Beach, Larry R.
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
```

```
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-14

Alignment Scores:
Pred. No.: 1,55e-218 Length: 3546
Score: 1822.00 Matches: 445
Percent Similarity: 40.86% Conservative: 29
Best Local Similarity: 38.36% Mismatches: 36
Query Match: 69.25% Indels: 652
DB: Gaps: 9

US-10-718-952-12 (1-510) x US-09-118-442-14 (1-3546)
QY 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyThrGluThrGluIle 20
Db 60 ATGTTTCATCGAGAGAGCTTCCGGCTGAGAGGCCCAAGCTGCGGTACAGGCGCCGAGATC 119
QY 21 GlnSerValTyraSnhTyGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyx 40
Db 120 GAGTCGAGAGTACCGGTACGACACACAGAGCTGTGACAGAGGCGCAAGAGAGCGCCCTCA 179
QY 41 GlnTrpIleValLysProLysSerValAsnTyGlnPheLysThrAsnThrHisValPro 60
Db 180 CGCTGGTGTGCTCGCGCCCAAGTCCGTCAAGTAACTTCGAGACAGAAACCGCGCTCCC 239
QY 61 LysLeuGly----- 63
Db 240 AAGCTCGGATATGTAAGGATGCAAGCGGCCCTAGCTCACTCTGTGAACCTCTCTCTC 299
QY 63 ----- 63
Db 300 CCGTGTCTGTCMAATCTCCGTGAGATCACTGTGCGGCTTCCCTTAAATCTCTTA 359
QY 63 ----- 63
Db 360 TGAATCTTACTGTGTTGCTGTAAGAGCAACCGTGTATTTGTGACAGCTACGACAC 419
QY 63 ----- 63
Db 420 ACTTGCCATCCGATGCGTCAAAATCAGCTGATTTGAATTCGATTCGATGCGCTT 479
QY 63 ----- 63
Db 480 TTCCATATTGATCATCTCCGCTACTGTGCAATGATTACAGAAACGTCCTTTTCTCT 539
QY 63 ----- 63
Db 540 TGAATTTGTCTTAGGCTTTTGTCTGTGACGATGAGCTGTATCAATTTGTTGATGTA 599
QY 63 ----- 63
Db 600 AGATCAATTCACGAGGAGGATGATGAGACAGACAGAACTCATTAACAGAAATTGA 659
QY 63 ----- 63
Db 660 TACTAGATTACTGAGCAGTGTGATGATGAGGCAATCTGCATCTGTGACCCCTTTGTT 719
QY 63 ----- 63
Db 720 TAATTCCTTCTTATGATGATGTGCTTACGTCAGCTGCTGTGTGTATGATGTGAGGC 779
QY 63 ----- 63
```

Db 780 TGTCAAGCCGCTTGTCTGTCTGCCAGGATGATGCCAACTTTTCTGTCTGTGTGTCAGG 839
Qy 64 ValMetLeuValGlyTTPGlyGlyAsnAnGlySerThrLeuThrGlyGlyValIleAla 83
Db 840 GTGAATGCTTGTGGGGTGGAGGACCAACAGGCTCCACCTGACGGCTGGGTCAATGGCC 899
Qy 84 AsnATGGLu----- 86
Db 900 AACAGGAGTGTAGTACTTAATTTGTCTATATTGCTTCCGTTGTTTTCAGTTATTA 959
Qy 87 -----AspIleSerTyrPalThr 92
Db 960 ATGGCCTAACAGAACTGAATTTTGTGTGGTTGTTCAAGGGATCTCATGGCCGACCC 1019
Qy 93 LysAspGlyIleGlnGlnIleAsnTyrPheGlySerLeuThrGlnIleSerAlaIleArg 112
Db 1020 AAGGACAGGTGTCCAGACCACTACTACGGCTC-CTCACCCAGGCTCCACCATCACA 1078
Qy 113 ValGlySerPheGlnGlnGlyGlnIleTyrAlaProPheLysSerLeuLeuProMet---- 131
Db 1079 GTGGGAGCTACACCGGGAGGAGATCTATGGCGCTCAAGAGCTCTCCCTCCATGTGA 1138
Qy 131 ----- 131
Db 1139 ATCTATATATAGCTTGACTAATACTCTCTTTTACTGAACCAACATACATACAAAG 1198
Qy 131 ----- 131
Db 1199 CATATTCCTGAAGTGTCTAGTTGATGTATTAATGAACCTGTCTTCAGGCCAGTGTTC 1258
Qy 131 ----- 131
Db 1259 TCAAGTAAACGGAAGTTAATCATTTGGGTTGAAAAAACAAAGTCTAATTTTGTGAAG 1318
Qy 131 ----- 131
Db 1319 GAAAGTTAATCTTAGCATATATGAAAAAGGGAGCACTGTAAAGAAAGTCTGAACAATC 1378
Qy 131 ----- 131
Db 1379 GACTCGTCTGCCAGTGTGTGAATCTTCACTTGCAAGTCAAAAGTTCGTGTGTTAGCCAA 1438
Qy 131 ----- 131
Db 1439 AGTTTCCAGACATCTTGATTAACCTGTGCAGTATTAAGATGTGTCTAATCTGTGCA 1498
Qy 131 ----- 131
Db 1439 GATTGCGACAGCTCGGTTTGTATCTTCTTTTCATGACCAAGTGTAAACGTGTTTCA 1558
Qy 132 -ValAsnProAspAspIleValPheGlyGlyTTPAspIleSerAsnMetAsnLeuAlaAs 151
Db 1559 GGTTGAACCCAGACGACATGTGTTCGAGGCTGGGACATTAGCAACATGAACTGGCCGA 1618
Qy 151 PalAsnMetAlaAsnAlaGlyValPheAspIleAspLeuGlnIleLeuArgProTyrMet 171
Db 1619 CTCCTGACAGCAGGCGCCAAAGTCTGTGATTTGACCTGCAGAAAGCACTCAGGCCCTCAAT 1678
Qy 171 GlnSerMetValProLeu-ProGlyIleTyrAspProAspPheIleAlaIleAsnGlnG 191
Db 1679 GGAAGTCCATGTGTCCCTTCCCGGATCTATGATTCGGAATTCATCGCGGCTTAACAGG 1738
Qy 191 LngLysArgAlaAsnAsnValIleLysGlyTThrLysGlnGlnIleValGlnIleIleL 211
Db 1739 GCTCTCGCGCCAAACAGTGCATCAAGGGACCAAGAAAGAACAGGGTGGAGCATCATCA 1798
Qy 211 YAspIle----- 213
Db 1799 AGGATATACGGTATATGATATGATGATCTAACGTCCTTGCTGAAGGTGACCCAGTG 1858
Qy 213 ----- 213
Db 1859 CAACCTAAACAAATTAATTAATACTATGAAATTTGGTAAATATACATACATACAGACAT 1918

Qy 213 ----- 213
Db 1919 AATGTTTAAACCGTGCACTTAGAGCTGTGATGTATGTGGACAATTTGACATTGCAATA 1978
Qy 213 ----- 213
Db 1979 TACAGTACCGCTCACTTGATGAGACTCCACAAAGAACTAAACTACTGAAGCTTAA 2038
Qy 213 ----- 213
Db 2039 GCAACTATTCGTAGCTAATGATGTAATTTGTGACATGTGTTGAAGATCTAGATTACGT 2098
Qy 213 ----- 213
Db 2099 GGTGAAGAAATATGCTTCACTAGTATATGATATCCATTACAGAAAGCAATGCTTATGTA 2158
Qy 214 -----LysAlaPheLysGlnAlaThrLysValAspLysValValIleLeuTyrPT 230
Db 2159 GCTAATGAACAGGAGTTTAGGGAGAAAGAACAAAGTGCACANAGATGTTGTGTGTGA 2218
Qy 220 hrAlaAsnThrGlnArgTyrSerAsnLeuValGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACACTGAAGATATAGCAATGTGTGCGTGTCTCAACGACAGATGAGAAATC 2278
Qy 250 euLeuAlaAlaValAspArgAsnGlnIleGlnIleSerProSerThrLeuTyrAlaIleA 270
Db 2279 TACTGGCATCTGTGACAAAGAACGAGCGAGATCTCACATCAACATATATGTCATTG 2338
Qy 270 laCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro- 289
Db 2339 CCTGTGTATGGAGGGGGTGCCTTCAATGAGAGCCCCAGAACACCTTTGTGCTTG 2398
Qy 289 ----- 289
Db 2399 GTGCGTGTGTGTGTGTGCAAAAGCCTCATGTGTGATTTCTGTCCAAAGTTTC 2458
Qy 290 -----GlyLeuIle 292
Db 2459 ATGTGTGTATTTCTGTCCAGGCTTATATACCTGTGACATGTCGTAGGGCTGATT 2518
Qy 292 AspLeuAlaIleAlaArgAsnThrLeuIleGlyIleAspAspPheLysSerGlyGlnThr 312
Db 2519 GATCTGTCTATAAAAACACTGCTTGATGTGTGTGATGACATCAAGTGTGACAGACC 2578
Qy 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLys----- 328
Db 2579 AAGATGAATCTGTCTGTGTGATTTCTGTGTGTGCTGGAAT-AAAGGTGGAACCTA 2637
Qy 328 ----- 328
Db 2638 GATCTCTCTTCTATTAAGATGAAGTGTTTTGGCAAAATGACCTTATTCATTAACCTC 2697
Qy 329 -----ProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 2698 TTCTATATTTTCATTTTCATGACGCCCACTTAATGTGAGTAAACCACTTGGGAAAC 2757
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db 2758 AACGATGGCATGAACCTGTCTGCCCTTCAACATTCAGGTCCAAAGAGATCTCCAAAGC 2817
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnHisPro 380
Db 2818 AACGTGTGTGATGACATGTGTCTCGAGCAATGCACTCTATGAGACCCGCGCAGCATCCC 2877
Qy 381 AspHisValValValIleLys----- 387
Db 2878 GATCATGTGTGTGTATCAAGGTCTGTAGCTGATCTTCACTCTGTTAAAGTTGACAT 2937
Qy 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGCAGATTACATGAAACTTGTCACCTCTTTGTGCAAGTATGTGCGGATACGTG 2997

QY 393 G1yAspserLySaRgAlMeRaspJUrThrSerGluLePheMeG1yG1yLysAsn 412
Db 2998 GGAGACAGAAAGGGCTATGAGCAGATACCTCAGAGATCTTCATGGCGGCAAGAAC 3057
QY 413 ThrIleValLeuH1aSnThrCySglUAspSerLeuLeuAla1ProIle1LeuAsp 432
Db 3058 AACCATGGCTGACAAACCTGTAGAGCACTGCTCCGCGCACTATCATCTTGAT 3117
QY 433 LeuValLeuLeuAla1LeuSerThrArgIleGlnPheLysAla1GluAsn----- 449
Db 3118 CTGGTCTCTTGCTGAGCTCAGCAGAGATCCAGTGAAGCTGAGGAGGGGTAAGA 3177
QY 449 ----- 449
Db 3178 GCGCCCAAGTATTAACTGAAGACGCTGACAGTATAGTATAGACTTTTAAATA 3237
QY 450 -----GluG1yLysPheH1aSerPheH1aProValAlaThrIle 462
Db 3238 CCTTCTGTGCTCTTATGAGCAAAATTCACCTCTCCACCGGTGCGCACATC 3297
QY 463 LeuSerThrLeuThrLysAlaProLeu----- 471
Db 3298 CTGAGCTACTCACCAGGACCCCTGTAAGCTTTTCTCTGCAATCCGCGCATCACTG 3357
QY 472 -----Val 472
Db 3358 CACTGGTTTGCTTCATCCAGCACTGATGCTCTTGAACCTGAACAGAGTT 3417
QY 473 PROProG1yThrProValAlaSnAlaLeuSerLySglUArgAlaMeLeuGluAsnIle 492
Db 3418 CCCCCGACACCGGTGTGAACCTCTGCGCAAGACAGACGCGCATGCTGAGAAACATC 3477
QY 493 MetArgAlaCySValG1yLeuAlaProGluAsnAlaMeIleLeuGluThrLys 510
Db 3478 ATGAGGCGCTGCTGGCTGGCGCCCAAGAAACAATGATCTGAGTACAG 3531

RESULT 7
US-09-677-064-14

Sequence 14, Application US/09677064
Patent No. 6291224
GENERAL INFORMATION:
APPLICANT: Martino-Calc, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
TITLE OF INVENTION: Genes Controlling Phyrate Metabolism in
FILE REFERENCE: 0706D
CURRENT APPLICATION NUMBER: US/09/677,064
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/055,446
PRIOR FILING DATE: 1997-08-11
PRIOR APPLICATION NUMBER: 60/055,526
PRIOR FILING DATE: 1997-08-08
PRIOR APPLICATION NUMBER: 60/053,944
PRIOR FILING DATE: 1997-07-28
PRIOR APPLICATION NUMBER: 09/118,442
PRIOR FILING DATE: 1998-07-17
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 3546
TYPE: DNA
ORGANISM: Zea mays
US-09-677-064-14

Alignment Scores:

Pred. No.: 1.55e-218 Length: 3546
Score: 1822.00 Matches: 445
Percent Similarity: 40.86% Conservative: 29
Best Local Similarity: 38.36% Mismatches: 36
Query Match: 69.25% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-12 (1-510) x US-09-677-064-14 (1-3546)
QY 1 MetPheIleGluAsnPheLysValG1ySerProAsnValLysThrGluThrGluIle 20
Db 60 ATGTTTCAATCGAGAGCTTCCGGCTGAGAGCCCACTGCGGTACGGCCGACGAGATC 119
QY 21 GlnSerValThrAsnThrGluThrGluLeuValH1aGluAsnArgAsnGlyThrTyx 40
Db 120 GAGTCGAGATACCGGTACAGACAGACGAGCTGTATCAGAGGCGCAAGAGACGGCGCTCA 179
QY 41 GlnThrIleValLysProLysSerValAsnThrGlnPheLysThrAsnThrH1aValPro 60
Db 180 CGTGGGTGCTGCCGCCAAGTCCGTCAAGTAACTTCGAGCCAGAACCGCGTCC 239
QY 61 LysLeuGly----- 63
Db 240 AAGCTCGGTATGTAAGGATGAGGAGCGGCTAGCCTCACTCTGTGAACCTCTCTC 299
QY 63 ----- 63
Db 300 CGGTGCTAGTCAAAATCCTCGTGAGATCAACTGTGCGGGTCCCTCTAAATCTTAA 359
QY 63 ----- 63
Db 360 TGAATCTTACTGCTTTCCTGAAGACAAACGTGTATTTGTGACAGTACGACAC 419
QY 63 ----- 63
Db 420 ACTTGCCATCCGATGCGTCAAAATCAGCTGATTTGAATTCGATTGATGTCCTT 479
QY 63 ----- 63
Db 480 TTCCATATTGAGATCATCCTCGCTACTGTGCAATGATTAAGAAAAGTCTTTTCTC 539
QY 63 ----- 63
Db 540 TGAATTTGCTTATGAGCTTTTGTCTGTGACGTGAGCTGTATTCATTTTTCANGTA 599
QY 63 ----- 63
Db 600 AGATCAAAATTCAGACAGGAGAGATGACAGACAGACAACTATTACTAGCAATTGA 659
QY 63 ----- 63
Db 660 TACTAGATTAOTGGCAAGTGAGATACGGGCAATTCGCATCTGAGACCCCTTGT 719
QY 63 ----- 63
Db 720 TAATTCCTTCTATGATGATGTGCTTACGTGACGCTGTGTGTATGATGTGATGAGC 779
QY 63 ----- 63
Db 780 TGTCAAGCCGCTTCTCTGTCCAGCGATGATGCCAATTTTCTGTGTGATGTCAGG 839
QY 64 ValMetLeuValG1yTrpG1yG1yAsnAsnLysSerThrLeuThrG1yValIleAla 83
Db 840 GTGATGCTTGTGGGTGGGAGGAGCAACAGGGTCAAGCTGAGCGGTGATTC 899
QY 84 AsnArgGlu----- 86
Db 900 AACAGGAGTGAATGATTAATTTGTCTATATTCCTTCCGTGTTCAGATTATTA 959
QY 87 -----AspIleSerThrAlaThr 92
Db 960 ATGGCTTAACAGAACTGAATTTTGTGTGTGTTTTCAGGGGATTCATGCGGAC 1019
QY 93 LysAspLysIleGlnAlaAsnThrPheG1ySerLeuThrGlnAlaSerAlaIleArg 112
Db 1020 AAGGACAAAGTGCAGACCAACTATCAGGCTC-CTCACCAAGGCTTCACATCA 1078
QY 113 ValG1ySerPheGlnG1yGluGluIleThrAlaProPheLysSerLeuLeuProMet--- 131
Db 1079 GTCCGCACTCAACGGGAGAGATCTATGCGCGTTCAAGAGCCTCTTCCCATGTGTA 1138


```

QY 131 ----- 131
Db 1139 ATCTATTATAGACTTGACTTAATCTCTCTTTTACTGAAACCAACATATACAAAG 1198
QY 131 ----- 131
Db 1199 CATATTCCGTAAAGTCTAGTTGATGTATATAAATGAACCTGTCTTTCAGGCCAGTGTGC 1258
QY 131 ----- 131
Db 1259 TCAAGTAAACGAAGTATATCATTTGGTTGAAAAAACAAGGTCTAAATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAAAGTTAACTTAGCATATAGAAAGGAGAGCACTGTAAAGAGTGTGAAACATC 1378
QY 131 ----- 131
Db 1379 GACTCGGTCTGCAGTGTGTGATCTTCTGCAAGTCAAAAAGTTCTGTGTAGCCCA 1438
QY 131 ----- 131
Db 1439 AGGTTCCAGCATCTTGTGATTAACACTGTGAGATTGACAGATGTGTACTGTGTGA 1498
QY 131 ----- 131
Db 1499 GATTCGCAGACTCGGTGTTGTATTCTCTTTTCATGACCAAGTGTAAACGTGTTTCA 1558
QY 132 -ValAsnProAspAspIleValPheGlyValTyrPheAsnMetAsnLeuAlaAs 151
Db 1559 GGTTGAACCCAGACGACATTTGTGTGGAGGCTGGACATTTAGCACAATAACTTGCCCA 1618
QY 151 PAlaMetAlaArgAlaValPheAspIleAspLeuGlnValGlnLeuArgProTyrMet 171
Db 1619 CTCCTATGACCAAGGCGCAAGTGTGTGATATTGACCTGCAGAAAGCACTGAGCCCTACAT 1678
QY 171 LeuIleuSerMetValProLeu-ProGlyIleTyrAspProAspPheIleAlaAlaAsnGlnG 191
Db 1679 GGAAGTCCAGTGGTCCACTTCCCGGTATGTATGATCCGACCTTCCGGGCTAACCAAG 1738
QY 191 LuGIuArgAlaAsnAsnValIleLeuGlyThrLeuGlnGluGlnValGlnGlnIleLeu 211
Db 1739 GCTTCGCGCCCAACAGTGTATCATCAAGGAGCAACAAGAAAGACGGTGGAGCATATCA 1798
QY 211 YAspIle----- 213
Db 1799 AGGATATACGATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 1858
QY 213 ----- 213
Db 1859 CAACCTAAAACAATATAATACTACTATGAAATTTGGTAATATATACATATACAGACAT 1918
QY 213 ----- 213
Db 1919 ATTGTTTAACCGGTGACTTAGAGTCTGCAATGATGTTTGGACAATTTGACATTGCATA 1978
QY 213 ----- 213
Db 1979 TACAGTGAACCGCTCACTTGTCATGAGAGACTCAACAAGAACTAAAACTACTGAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTCTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2098
QY 213 ----- 213
Db 2099 GGTGAAGAAATATGCTTCACTAGTATATAGTAAATCATTTACGAAGCAATGCTTATGTA 2158
QY 214 -----LysAlaPheLeuGlyValAlaThrLeuValAspLeuValValLeuTyrP 230
Db 2159 GCTAATGAACACGAGGAGTTAGGAGAGAAAGAACAAAGTGGACAAAGATAGTGTGTGTGA 2218

```

```

QY 230 hrAlaAsnThrGluArgTyrSerAsnLeuValGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAACACACTAAAGATATAGTATGTCGCTGTGTCTCAACGACAGATGAGAAATC 2278
QY 250 euLeuAlaIleValAspArgAsnGluAlaGluIleSerProSerThrLeuTyrAlaIleA 270
Db 2279 TACTGCACTCTGTGACAAAGAACGAGCGAGATACCATCAACACATATATGCAATTG 2338
QY 270 IeCyValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro 289
Db 2339 CCTGTGTATGAGAGGGGCGCTTCATCATATGAGAGCCCGACAGAACACTTGTGTGCTG 2398
QY 289 ----- 289
Db 2399 GTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2458
QY 290 -----GlyLeuIle 292
Db 2459 ATGTGTGTATTTCTGTCCAAAGCTTATTTACTGTGTGATGTGTGTGTGTGTGTGT 2518
QY 293 AspLeuAlaIleAlaArgAsnThrLeuIleGlyIleAspAspPheIleValGlnThr 312
Db 2519 GATCTTGTCTATAAACAACACTGCTTGTGATGTGTGTGACGACTTCAAGATGGAACAGAC 2578
QY 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLeu----- 328
Db 2579 AAGATGAATCTGTCTGTGTGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2637
QY 328 ----- 328
Db 2638 GTATCTCTCTTATTAAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2697
QY 329 -----ProThrSerIleValSerTyrAsnIleGlnIleValAsn 340
Db 2698 TTCTATTATTTCTATTTTCATGACGCCCACTCAATCGTAGCTAACACACTTGGGAAC 2757
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLeuGluIleSerLysSer 360
Db 2758 AACGATGCAATGAACTGTCTGTGCTTCCCTTCAMAACTTCAGTCCAGAGATCTCCAAAGAC 2817
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyValIleIlePro 380
Db 2818 AACGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2877
QY 381 AspHisValIleValIleLeu----- 387
Db 2878 GATCATATCGTTGTATCATCAAGGTCTGTAGCTGATCTTTCACCTCGTTAAAGTTGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGCAGATTATTAATTGAACCTTGTCACTTTTGTGTGCAATATGTGCGGTACG 2997
QY 393 GlyAspSerLysArgAlaMetAspGluTyrThrSerGluIlePheMetGlyValLysAsn 412
Db 2998 GAGACACAGCAAGGGCTATGACAGATACACTCAGAGATCTTATATGGCGGCAAGAAC 3057
QY 413 ThrIleValLeuHisAsnThrCysGluAspSerLeuLeuAlaIleProIleIleLeuAsp 432
Db 3058 ACCATCGTGTGACCAACACCTGTAGAGACTCGTCTCGCGGCACTATCATCTTGTAT 3117
QY 433 LeuValIleLeuValIleGluLeuSerThrArgIleGlnPheValIleGluAsn----- 449
Db 3118 CTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3177
QY 449 ----- 449
Db 3178 GCCCCCAAGTATTAACCTGAAGACGCTGCAGCTAGTATATGACCTTTTATA 3237
QY 450 -----GluGlyLysPheHisSerPheHisProValAlaThrIle 462
Db 3238 CCTGTGTGTCTCTTATGACAGAGAAATTCACCTCTTCCACCGGTGGCCACATTC 3297
QY 463 LeuSerTyrLeuThrValProLeu----- 471

```

Db 3298 CTGACACTCTACCAAGGCAACCCGGGAAGCCTTTCTCTGATCCGGGCATCAGTG 3357
QY 472 -----Val 472
Db 3358 CACTGCGTTTGCTTCAATCCAGCACTGATGCTCTCTTGAAA CTTGAACAAGGTT 3417
QY 473 PROPGIYTHrProValAlaAsnAlaLeuSerLysGlnArgAlaMetLeuGlnAsnIle 492
Db 3418 CCCCCTGGACACCGGTGTGTAAACCTCTGGCCAAAGCAAGCGCAATGCTGAGAAACATC 3477
QY 493 MetArgAlaCyValGlyLeuAlaProGlnAsnAsnMetIleLeuGlnTyrllys 510
Db 3478 ATGAGGCGCTGCGCTTGCGGTGCGCCCAAGAACATGATCTCTGAGTAACAAG 3531
RESULT 8
US-09-734-237B-72
; Sequence 72, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 72
; LENGTH: 1602
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-72
Alignment Scores:
Pred. No.: 1,42e-164 Length: 1602
Score: 1389.50 Matches: 278
Percent Similarity: 69.66% Conservative: 87
Best Local Similarity: 53.05% Mismatches: 138
Query Match: 52.81% Indels: 21
DB: Gaps: 8
US-10-718-952-12 (1-510) x US-09-734-237B-72 (1-1602)
QY 3 IlegluAsnPhelyValGluSerProAsnValLysTyrlThrGluThrGluIleGlnSer 22
Db 25 ATCACCCTCCGTTAAAGTAGTACCGACAGATGCAAGTCAAGAGCAACAAGCGCTCAGC 84
QY 23 ValTyrlAsnTyrlGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyrlGlnTrp 42
Db 85 AAGTACAGCTACGAAAGTCT--GTAGTTACGAAGACAGCTAGTGGCCGCTTCAT-- 138
QY 43 IleValLysProLysSerValAsnTyrlGlnPhelyThrAsnThrHisValPro--Lys 61
Db 139 ---GTAAACCCCACTTCAAGACTACGCTCAAACTGACTTGAAGAAAGCCGGAAGAA 195
QY 62 IeuglyValMetIleuValGlyTyrlGlyLysAsnAsnGlySerThrLeuThrGlyVal 81
Db 196 CTAGAGAAATATGCTCATGTGGTTAGGTGGCAACAAATGAGCTCACTTATGCGCTCGTA 255
QY 82 IleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsnTyrl 101
Db 256 TTGGCGCAATTAAGACATGTGAGTTCAACTAAGAGCGCTTAAGCAACCAATCTAC 315
QY 102 PheGlySerLeuThrGlnAlaSerAlaIleArgValGly---SerPheGlnGlyGlnGlu 120
Db 316 TTGGGCTCATATGACTTAAGTCTTACCTTGAAGACTGGATATGATCGGAGGGGAATGAC 375
QY 121 IleTyrlAlaProPhelysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 376 GTTATATGCTCTTTAACTCTCTGTGGCCATGGTTAAGCCCAAGCACTTGTGCTCT 435
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaIleArgAlaLysValPheAsp 160
Db 436 GGTGGGACATCATATACGACGATCTTATACAGACTATGACAGAAAGTCAAGTCTCGAA 495
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrlMetGluSerMetValProLeuProGlyIle 180
Db 496 TATGATCTGCAACAACGCTTGAAGCGCAAGATGCTTGAGGAAGCCTTCTCTTCATTT 555
QY 181 TyrlAspProAspPheIleAlaAlaAsnGlnGluValGlnAlaAsnValIle----- 198
Db 556 TACTAACCCTGATTTCAATTCAGCTTAATCAATAGAGAGCCCAATATGATGATCAATTTG 615
QY 199 -----LysGlyThr-----LysGlnGlnGlnValGlnGlnIleIleLys 211
Db 616 GATGAAAAGGCAACGTAAACCAAGAGGGTAAAGTGAAGCCATCTGCAACGATCAGAGCC 675
QY 212 AspIleLysAlaPheLysGlnAlaThrLysValAspLysValValIleuTrpThrAla 231
Db 676 GATATCCAGAAATTTCAAGAAAGAAAAGCCCTTGATTAAGTAATGCTTTGGACTGCA 735
QY 232 AsnThrGluArgTyrlSerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeu 251
Db 736 AATACTGAGAGTACGTAAAGTATCTCTGCTGTGTAAATGACACCAAGAAAACCTCTTG 795
QY 252 AlaAlaValAspArgAsnGluAlaGluIleSerProSerThrLeuTyrlAlaIleLys 271
Db 796 CAGTCTATTAAGATGACATGAGCAAGAGTTGCTCTTCCAGATCTTGGACAGCAACT 855
QY 272 ValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeu 291
Db 856 ATCTTGAAGGTGTCCTTATATATGTTTCAACCGCAATATCTTTGTTCCCGGCTTG 915
QY 292 IleAspLeuAlaIleAlaArgAsnThrLeuIleGlyLysAspAspPheLysSerGlyGln 311
Db 916 GTTCAGCTGTCGACATGAGGATCATTCATTCGGGAGAGATCTCAAGTGGGAGAA 975
QY 312 ThrLysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLysProThrSer 331
Db 976 ACCAAGTTGAAGTCTGTTCTGGCCAGTCTTATGATGATGACAGATTAACCGGCTCTCC 1035
QY 332 IleLysSerTyrlAsnHisLeuGlyLysAsnAspGlyMetAsnLeuSerAlaProGlnThr 351
Db 1036 ATTCATCTCTTAACCATTTAGGCAATATGACGCGTTATCTTATCTGCTCAAAACAA 1095
QY 352 PheArgSerLysGlnIleSerLysSerAsnValValAspAspMetValAsnSerAsnAla 371
Db 1096 TTTTAGTCTTAAGAGATTTCCAAAAGTTCTGTCAATGATGACATCATGCGCTTAATGAT 1155
QY 372 IleLeuTyrlGluPro-----GlyGluHisProAspHisValValIleLysTyrlVal 389
Db 1156 ATCTTGTACATGATTAACATGCGTAAAGAAAGTGAACCTGATGTGATCAATTAATG 1215
QY 380 ProTyrlValGlyAspSerLysArgAlaMetAspGlyTyrlThrSerGluIlePheMetGly 409
Db 1216 AAGCCCGTCGGGACTCAAAAAGTGCATGACGAGTGTTCAGAGATGTGATGTAGGT 1275
QY 410 GlyLysAsnThrIleValLeuHisAsnThrCysGluAspSerLeuLeuAlaAlaProIle 429
Db 1276 GGCCATAACCGGATTTCCATTCACATGTTCGCAAGATTTCTTATCGGCTACGCGCTTG 1335
QY 430 IleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGlnPhelys----- 446
Db 1336 ATCATGATCTTTTATGATGATGATGATGATTTGTACAAAGTCTCTTAAGAAAGTGAAC 1395
QY 447 -----AlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSer 464
Db 1396 CCAAGTTAAAGAAAGATGCTGGGAAATTCGAGAACTTTATCCAGTTTAACTTTGAGT 1455
QY 465 TyrlLeuThrLysAlaProLeuValProProGlyThrProValValAsnAlaLeuSerLys 484
Db 1456 TACTGTATAAAGCTTCATTAACAAGACCAAGATTTCAACCGGTGAATGGCTTAACAAG 1515
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

QY 485 GlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluAsn 504
DB 1516 CAAGAAGCCGCTTGAAGAAATTTTAAAGATGTTGATTCGATTCCTTCAAAACGAA 1575
QY 505 MetIleLeuGlu 508
DB 1576 CTAAAGATTCGAA 1587

RESULT 9
US-09-734-237B-74
; Sequence 74, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734, 237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494, 921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic gene derived from Saccharomyces cerevisiae myo-inositol
; OTHER INFORMATION: -1-phosphoric synthase, having numerous codons replaced with other
; OTHER INFORMATION: s encoding the same amino acids to reduce free energy of folding,
US-09-734-237B-74

Alignment Scores:
Pred. No.: 1,429-164 Length: 1605
Score: 1389.50 Matches: 278
Percent Similarity: 69.66% Conservative: 87
Best Local Similarity: 53.05% Mismatches: 138
Query Match: 52.81% Indels: 21
DB: 4 Gaps: 8

US-10-718-952-12 (1-510) x US-09-734-237B-74 (1-1605)

QY 3 IlegluAsnPhelValGluSerProAsnValIysTyrThrGluThrGluIleGlnSer 22
DB 28 ATCATCTTCTGTTAAAGTTGTAAGTGAACAATGTAATCAAGAACAGAACGACGCGTACT 87
QY 23 ValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyrGlnTrp 42
DB 88 AAATCTCTTACGAAACCGCT--GTAAGTAACTAAACGCTTCGTCGTTTCGAT-- 141
QY 43 IleValIysProIysSerValAsnTyrGlnPhelYsrThrAsnThrHisValPro---Lys 61
DB 142 ---GTTATCCCACTGTTCAAGACTACGATTCACAACTGCGATCGAAGAACGGGAAAG 198
QY 62 LeuGlyValMetLeuValGlyTyrGlyIysAsnAsnGlySerThrLeuThrGlyGlyVal 81
DB 199 CTGGGTATCATCATGCTGAGCGCTGGGTGTAAACAACGCTCATCTGTTGTCATCTGTT 258
QY 82 IleAlaAsnArgGluAspIleSerTrpAlaThrIysAspIysIleGlnGlnAlaAsnTyr 101
DB 259 CTGGCAACCAACACACACTAGAAATTCACAGCTAAAGAGGTGTTAAACAGCCGAACTAC 318
QY 102 PheGlySerLeuThrGlnAlaSerAlaIleArgValGly---SerPheGlnGlyGluGlu 120
DB 319 TTGATTTTATATACACAGTCTTACTCTGAAGCTGGGCGATTGATGCTGAAGGTAAACGAC 378
QY 121 IleTyrAlaProPheIysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 379 GTTACGCTCGGTTCAACTCTCTGCTGGCGGATGATGATCCGAACGACCTTGTGTTCT 438

QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIysValPheAsp 160
DB 439 GGTGGATATCAACAACCGCATCTGTACAGCAATACAGCGTTCTTCAAGTTCTTCGAA 498
QY 161 IleAspLeuGlnIysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 499 TATGATCTGCAACAGCGTCTGAAGGCTAAGATGTCCTCTGTTAAACCACTGCGCTCATC 558
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluGluValIleAsnValIle----- 198
DB 559 TACTACCCGGAATTTATCGACACTAACAGCAAGCAAGTGTCAACAGTATCAACCTG 618
QY 199 -----LysGlyThr-----LysGlnGluGlnValGlnIleIleIys 211
DB 619 GACGAAAGGGTAAACGTTATCAACCGGTGTATGAGACTCACCTGAGGATACCGTCT 678
QY 212 AspIleIysAlaPheIysGluValThrIysValAspIysValValIleuThrAla 231
DB 679 GATATCCAGAACTTCAAGAGGAAACGACCTGACAAAGTATGCTGACTGACTGCT 738
QY 232 AsnThrGluArgTyrSerAsnLeuValIleGlyLeuAsnAspThrMetGluAsnLeu 251
DB 739 AACACTGAACGTTTACGTAAAGTATCCCGGCTGTAAACGATATGAGAAACCTGCTG 798
QY 252 AlaAlaValAspArgAsnGluAlaGluIleSerProSerThrLeuTyrAlaIleAlaCys 271
DB 799 CAATCTATCAAGACACGACGAGAAATCGTCCGTCACCATCTCTCTGATCT 858
QY 272 ValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeu 291
DB 859 ATCTCGAAGAGGTTACCGTATCAACGCTCTCGGCAACACATTGCTACCGGCTG 918
QY 292 IleAspLeuAlaIleAlaArgAsnThrLeuIleGlyIysAspAspPheIysSerGlyGln 311
DB 919 GTACAGCTGGCTGAACAGCAAGGTACTTCACTGCTGTGACGATCGAAATCTGCGCCAG 978
QY 312 ThrIysMetIysSerValLeuValAspPheLeuValGlyAlaGlyIleIysProThrSer 331
DB 979 ACTAAATCGAAATCTGTACTGCGACAGATTCCTGTTGACCGCTGTATCAACCGGTTCT 1038
QY 332 IleValSerTyrAsnHisIleGluGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThr 351
DB 1039 ATCGCTTCTTATACCACTGCTGCTGTAACAACGCTACACCTTCTGCTCCGAACG 1098
QY 352 PheArgSerIysGluIleSerIysSerAsnValValAspAspMetValAsnSerAsnAla 371
DB 1099 TTCGTTCTTAAAGAAATCTTAAATCTCTGTATCGACGACATCATGCGCTTCAACGAC 1158
QY 372 IleLeuTyrGluPro-----GlyGluHisProAspHisValValIleIysTyrVal 389
DB 1159 ATCCGTGTACACGACCAACCTGGTAAAGATGATCACTGTATGTTATCAAAATACATG 1218
QY 390 ProTyrValIleGlyAspSerIysArgAlaMetAspGluTyrThrSerGluIlePheMetGly 409
DB 1219 AAACCGTTGGATTTCTTAACTTCTTATGACGAATCTACTCTGAACCTGATGCTGGCC 1278
QY 410 GlyIysAsnThrIleValIleuHisAsnThrCysGluAspSerLeuLeuAlaAlaProIle 429
DB 1279 GGTACCAACCGATATCTATACCAACAGTTGTGAAGACTGCTGCTGCTGCTACCCGCTG 1338
QY 430 IleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGlnPheIys----- 446
DB 1339 ATCATCGACCTGCTGTATATACCTAAATTCGTATCCCGGTATCTTAAACAAAGTTGAC 1398
QY 447 -----AlaGluAsnGluGlyIysPheHisSerPheHisProValAlaThrIleLeuSer 464
DB 1399 CCGGTTAAAGAGAGCTGGCAATTCGAAACATTTACACCGGTTCTACCTCTGCTCT 1458
QY 465 TyrLeuThrIysValaProLeuValaProProGlyIThrProValIleAsnAlaLeuSerIys 484
DB 1459 TACTGGCTGAAGAGCTCGCTGACTGTCACAGGCTTCCACCGGTTAAACGCTGAACAA 1518

QY 485 GlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluAsnAsn 504
Db 1519 CAGCGTACCGCTCTGGAAACCTTCTCGCTGCTGATCGCGCTCCGACCAAGAA 1578
QY 505 MetIleLeuGlu 508
Db 1579 CTGCGTTTCGAA 1590
RESULT 10
US-09-248-796A-3131
; Sequence 3131, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3131
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3131
Alignment Scores:
Pred. No.: 5.9e-164 Length: 1578
Score: 1384.50 Matches: 266
Percent Similarity: 71.29% Conservative: 99
Best Local Similarity: 51.95% Mismatches: 132
Query Match: 52.62% Indels: 15
Gaps: 5
US-10-718-952-12 (1-510) x US-09-248-796A-3131 (1-1578)
QY 9 GluSerProAsnValIleYThrThrGluThrGluIleGlnSerValIleYAsnIleYThr 28
Db 37 AAATGCTCCAAATCTGTCACTAAGACGATCATTTATATACCAATTAATTACTATATAAAC 96
QY 29 ThrGluLeuValHisGluAsnArgAsnGlyThrIleYThrGluIleValIleYAsnProIleSer 48
Db 97 TCT--GTTGTTGAAAGATGCTTAATGCT--AAATTCATGTTACACCACTGCT 147
QY 49 ValAsnIleYThrGluPheIleYThrAsnThrHisValProIleYAsnGlyValMetLeuValGly 68
Db 148 AGTACCTATGATTCGAAAGTTGATTTAAAGTCCCTAAGGTTGTTATTTGTTAGTCCGT 207
QY 69 TrpGlyGlyAsnAsnGlySerThrIleYThrGlyValIleAlaAsnArgGluAspIle 88
Db 208 ATGGCGGTAACAAAGCTACTACATTTAGGTGACCAATTGACAGACAAACACAAATATT 267
QY 89 SerTrpAlaThrIleYAspIleYThrGlnIleAlaAsnIleYThrPheGlySerLeuThrGlnIle 108
Db 268 TCTTTGAAACAAAGAGGTGTGTTAAACCAAAATTACTATGTTCTGTTACTGAACGA 327
QY 109 SerAlaIleArgValGly-----SerPheGlnGlyGluIleYThrAlaProPheIle 126
Db 328 TCCACTGTCAAATAGGTGTGATTAAGAACTGATGAGGATTTACGTTCCATTTAAC 387
QY 127 SerLeuLeuProMetValAsnProAspAspIleValPheGlyGlyTrpAspIleSerAsn 146
Db 388 TCCATTTGTTCCAAATGTTAAACCAATGATTTGTTGATGAGGTGGATTTAGTGT 447
QY 147 MetAsnLeuAlaAspAlaMetAlaArgAlaIleYAsnValPheAspIleAspLeuGlnIleGln 166
Db 448 TTGCAATTAGACCAAGATTAAGAGCTAAAGTCTTAAGCTTGAATTAACAAAAGCAG 507
QY 167 LeuATGProIleYMetGluSerMetValProLeuProGlyIleYThrAspProAspPheIle 186

Db 508 TTGTATCTATCTTGTGGAGAACAGAAACCTTGGAAATCAATCTACTACCTCGATTTCACT 567
QY 187 AlaAlaAsnGlnGluIleArgAlaAsnValIleYAsn----- 199
Db 568 GCTTGAACCAAGTGAACGTCGACCAATGTTTAAACCAAGTCAATGCTGAAGTTTAA 627
QY 200 --GlyThrIleYAsnGlnGluIleValGlnIleIleYAspIleYThrAlaPheIleGlu 218
Db 628 ACTGATATATAAGTGGCGACGTTGAAATAATCAGAAAGATATCAGAGATTTCAAGGCC 687
QY 219 AlaThrIleYAspIleYValIleValLeuIleYThrAlaAsnThrGluIleYThrSerAsn 238
Db 688 AAAACGATTAAGTAAAGTTATTTTAAAGCTGCTATATCTGAAAGTAAAGTCTGAT 747
QY 239 LeuValIleGlyLeuAsnAspThrMetGluAsnLeuIleAlaValAlaAspArgAsnIle 258
Db 748 GTTTGCCAAATGTCATATGACACTGTCGACAACTTGATCAAGATATTAAAGATCACAC 807
QY 259 AlaGluIleSerProSerThrLeuIleYThrAlaIleAlaCysValMetGluAsnValProPhe 278
Db 808 GAAGAAATGCTCCAAAGTACTGCTTGGCGGTGCTCTTCTGAAATAAGTCCATAT 867
QY 279 IleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArg 298
Db 868 ATTAATGTTCCACCAACAAACACATTTGTTCCCGTGTATTAAGTAAAGTAAATAC 927
QY 299 AsnThrIleIleGlyIleYAspAspPheIleYSerGlyGlnThrIleYMetIleYSerValIleu 318
Db 928 GACTCATTCATTTGTTGATGATTTCAAGTCAAGTCAACCAAAATTAATCATCTGTTA 987
QY 319 ValAspPheLeuValIleYAlaGlyIleYAspProThrSerIleValSerIleYAsnIleu 338
Db 988 GCTCAATCTGTTGTCGATGCTGATCAACCACTTCTATTTGCTTATTAATCACTTG 1047
QY 339 GlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGluIleSer 358
Db 1048 GGTAAACAATGACGGTTCAATTTATCATCAACAAACAAATTTAATCTTAAGAAATTTCC 1107
QY 359 LysSerAsnValIleYAspAspMetValAsnSerAsnAlaIleYThr-----GluPro 376
Db 1108 AAACATCTGTTGTTGATGATTTATTAATTCACCAAGATTTATGACACACAGAAATCT 1167
QY 377 GlyGluHisProAspPheHisValIleYThrIleYThrValProIleYAspSerIleY 396
Db 1168 GGTGACAAAGTGAATCACTGATGCTGATTAATACCTTGCCAGCGTTGATTCATAA 1227
QY 397 ArgAlaMetAspGluIleYThrSerGluIlePheMetGlyGlyIleYAsnThrIleValIleu 416
Db 1228 GTTCCATGATGATGATCTATTCAAGAAATTAATGTTGGTGGTCCACAAATTAATATT 1287
QY 417 HisAsnThrCysGluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleu 436
Db 1288 CACAATGTTTGAAGATTCATCTGCTGTCACACATTTGATTTATGATGTTGGTGGCT 1347
QY 437 AlaGluLeuSerTrpArgIleGlnPheYAspIleGluAsnGlnGlyIleYAsnIleYThrPhe 456
Db 1348 ACTGAATTTGCAACTAAGATTCATCAAGTCCCTGTTAATGATGATGAAGAAATG 1407
QY 457 HisProValAlaThrIleLeuSerIleYThrIleYThrIleYAspIleYThrIleYThr 476
Db 1408 TATCCTGTTGATGATTTATGTTTACTGCTCAAGGCTCATTAAGCAACGACGAGATTC 1467
QY 477 ProValIleAlaAsnIleYSerIleYAspIleYAspIleYAspIleYAspIleYAspIleY 496
Db 1468 AAACCTATCAACGATTAACAAACCAAGTCAACAAATTAATGATCACTTTCAGTGTG 1527
QY 497 ValGlyLeuAlaProGluAsnAsnMetIleGlu 508
Db 1528 GTTGGTTTCAATTTGACACGAATTAAGATTTGAA 1563
RESULT 11
US-09-949-016-12608

Sequence 12608, Application US/09949016
Patent No. 581239
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12608
LENGTH: 77626
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) ... (77626)
OTHER INFORMATION: n = A, T, C or G
US-09-949-016-12608

Alignment Scores:

Pred. No.:	1,79e-141	Length:	77626
Score:	1229.50	Matches:	313
Percent Similarity:	41.22%	Conservative:	86
Best Local Similarity:	32.33%	Mismatches:	106
Query Match:	46.73%	Indels:	466
DB:	4	Gaps:	12

US-10-718-952-12 (1-510) x US-09-949-016-12608 (1-77626)

QY 6 PheLysValAlGusESePProhAmValLySryThrGluThrGluIleGlnSerValTyzAsn 25
DB 32903 TTCTTCGTGAGAGCCCGGACCTGGTCTACGCCCCGAGCCATCGAGCGCAATAACGAG 32962
QY 26 TyrGluThrThrGluLeuValHisGlu----- 34
DB 32963 TACCGAGAGACCGCGGTACGCCCGAGGGTGGCGTTCTCAAGCTGACCCAGGGGCTGAG 33022
QY 35 -----AsnArgAsnGlyThrTyrlnTrpIle 43
DB 33023 GAAGGAGCGTGTCTGGGAGAGCGGGCGGCTCGGGAGAGAGGACCGGAGATT 33082
QY 44 -----VallyPProlySserValAsnTyrlnPhelysThrAsn 56
DB 33083 CACTGCGCGCTGTCTTACAGGTGCACCCACGTCACGCGCTTCACCTTCCGAGACCGCC 33142
QY 57 ThrHisValPProlyLeuGlyValMetLeuValGlyTrpGlyValAsnAsnGlySerThr 76
DB 33143 CGGACGATGCCCCGCGCTCGGGGTCAAGCTTGTGGGTGGGGGAGAAACAAGCTTCACA 33202
QY 77 LeuThrGlyGlyValIleAlaAsnArgGluAspIleSerTrpAlaThrlyAsp----- 94
DB 33203 CTCAACGCGCGGTCTGCGCATGACGCGCTTGTCTGCGCCACCGCGACGCGCGCC 33262
QY 94 ----- 94
DB 33263 AAGGTGGGGCGGAGAGGGCGTGTGGTGAAGAGGCTTGTGGAGTCTTGGCGGCGG 33322
QY 94 ----- 94
DB 33323 GGGCGGAGCTTGCAGAGGGCGGGCTCTGCGGAGCAAGAGGGGCTGTAGGGGGTGGGCT 33382
QY 94 ----- 94
DB 33383 TCCGCGGGGCGAGGGGCTGCAAGGGCGAGGACCTGCCGGTCTGAGAGCCTGTAGTGG 33442

QY 94 ----- 94
DB 33443 GCGAGGACCTGACAGGGGCGAAGCTGTGTGGGCGGGAGGGCTTGCAGGGGCGGGGCC 33502
QY 94 ----- 94
DB 33503 TGGCTGAGGGGAGCCCGGTGAAGGTTTGGCGAGGTAGAGGGGACTCTCGGAGTCT 33562
QY 94 ----- 94
DB 33563 TACAGGGCGAGCTTAAGTGTCCGAGAGTCTCAGAGGAGGGGCTTGGGGGCGAGCGCTT 33622
QY 95 -----LysIle----- 96
DB 33623 CCGTCTGCTGTGAGCCTGGAATGAATAATGCCGAAGTGGGGCGGGGCTTGAAGT 33682
QY 96 ----- 96
DB 33683 CTAGGTGGGGCCAGCGCGGTGCGGCGCTTTCACCACTCCGACCTGCGCGT 33742
QY 97 -----GlnGlnAlaAsnTyrlPheGlySeryLeuThrGlnAlaSerAlaIleAsp 112
DB 33743 GTCTGCTGACCCCGAGAGGCCAATCTACGCTGCTGACTCAGCGGGGACCGTGA 33802
QY 112 gValGly---SerPheGlnGlyGluIleTyrlaPhePheLySeryLeuLeuProMe 131
DB 33803 CCTGGGCTTGAACCGCGAGGCGCAGAGGTGTTCGTACCTTTCAGCGGGTCTGCCAT 33862
QY 131 tValAsnProhAspAlaIleValPheGly----- 140
DB 33863 GGTGGCGCCCAAGACCTCTGTTCGA--TGTGGGCGGAGCCCTGGGCGGGGTGGGGCG 33921
QY 140 ----- 140
DB 33922 GGATGGAGATGGGTCTGAGAGGGCCAGATCCGGGCGGAGCCGAGTGTAGAGGTCCC 33981
QY 141 -----GlyTrpAspIleSeryAsnMetAsnLeuAlaAspAlaIleValPhe 159
DB 33982 CCCAGGCTGGAGCACTGTGTGCTGACCTGGCCAGCGAGCGGCGCGCAAGGTGT 34041
QY 159 eAspIleAspLeuGlnTyrlSeryGlnLeuArgProTyrlMetClySeryMetValProLeuProG 179
DB 34042 GAGCTGGGGGCTGCGAGAGCACTGTGGCCGACATGAGAGGCGCTCGGGCCCTTC 34101
QY 179 yLleTyrlAspProhAspPheIleAlaIleAsnGlnGluIleValAlaAsnAsnValIle 199
DB 34102 TGTTCATATCCCGAATTCATGTGCGGCAACAGAGCGCGCGGCAACATCATCCC 34161
QY 199 sGlyTyrlSeryGlnGluIleVal----- 206
DB 34162 AGGCTGGGTGCGAGAGGTGTGTCCATCCCGCATCCCTGTCTCTGCCACTGT 34221
QY 206 ----- 206
DB 34222 CTTCGCACTGTGTTCTCTTTTGGCCCAACCCCGCAATCCAGGCTCACTTGAACAC 34281
QY 206 ----- 206
DB 34282 CGGGAACTGAGTAGCCAGGCGTGTGAGTCCGAGAGTATGTACAACAACTTCCCTA 34341
QY 207 -----GlnGlnIleIle 211
DB 34342 CACCTCTCAATCTTCCCAACCTCTCAATCCAACCAACAGCTGAGAGCATCCGAG 34401
QY 211 eAspIleLeuValaPheLySeryGlnAlaIleThryValaAspLyValaIleValaLeuTrpThrAl 231
DB 34402 GGAATTCGAGACTTCCTCGCTAGCGCGGGCTGGAACAAGTCAATAGTCTGTGAGAGCG 34461
QY 231 aAsnThrGluArgTyrlSeryAsnLeuValaGlyLeuAsnAspThrMetGluAsnLeu 251
DB 34462 GAAACGAGAGGCTTCTGTGAGGATTCAGAGGCTCAACAGACACAGCGGAACCTGCT 34521
QY 251 uAlaAlaValaAspArgAsnGlu-Ala----- 259


```

QY 336 YLELYSProthrSerIleValSerTyrAsnHisLeuGIYAAsnAsnAspGlyMetAsn-L 346
DB 422 CCTAAGACCATGTCATGAGTATGACCAACCTGGGCAACAGATGGGAGAACCC 481
QY 346 euserAlaProGlnThrPheArgSerTyrGluIleSerTyrSerAsnValValAspAspM 366
DB 482 TATGGGCCCATTTGCAAGTTCGGCTCTAAGAGAGGTGTCCAAAGCAACGTGTGAGACTGGAGCA 541
QY 366 etValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisProAspHisValValValI 386
DB 542 TGGTCAGAGCAACCCAGAGCTCTATACGCCCGGCGAAGAGCTGACCTCGGTGCTCA 601
QY 366 IeLYTYrValProTyrValGIYAAspSerTyrAlaMetAspGlnTyrThrSerGluI 406
DB 602 TCMAAGTATGTCGCTGACGAGTGGGTGACAGACGGCGCTGGATGATATACCTGGAGAC 661
QY 406 IePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCysGluAspSerLeuLeuA 426
DB 662 TGATGCTGGCGGCAACCAACACACTGGTCTGCACAAACGTGTAGAGACTGGCTGCTGG 721
QY 426 IAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGlnPheL 446
DB 722 CCGCACCCATCATGCTGGACCTAGCGCTGCTGACCGAGCTGTGCCAGCGCGTGAAGCTTCT 781
QY 446 YSAlaGluAsnGluIleLysPheHisSerPheHisProValAlaThrIleLeuSerTyrL 466
DB 782 GCACGTGACATGACCCCGACGCGGACACCTTCACCCCGCTGCTGCCCTCACCTTCC 841
QY 466 eUTHrLYAlaProLeuValIProProGlyLThrProValValAsnAlaLeuSerLysGlnA 486
DB 842 TCTTCAAGGCGCCACTAGTGGCGCCCGGACGCCGGGTGTCAATGGCTTTCCGCCAGC 901
QY 486 rGAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluAsnAspMetI 506
DB 902 GCAGCTGATCGAAGAACATCTCAAGGCGCTGCGTGGGCTCCCGCACAGAACCAACATCC 961
QY 506 IeLeuGluTyrLys 510
DB 962 TCTGTGAACACAA 975

```

RESULT 13

```

US-09-902-540-1257/C
; Sequence 1257, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1257
; LENGTH: 34316
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1257

```

Alignment Scores:

```

Pred. No.: 1.02e-53 Length: 34316
Score: 529.00 Matches: 141
Percent Similarity: 46.92% Conservative: 103
Best Local Similarity: 27.12% Mismatches: 206
Query Match: 20.11% Indels: 71
DB: 4 Gaps: 13

```

US-10-718-952-12 (1-510) x US-09-902-540-1257 (1-34316)

```

QY 17 GIUTRGIUllEGInSerValTyrAsnTyrGIUTRGIUleu-----Val 32
DB 20023 GAAGTTGAAGTATGATGAACGCTTTAGATATCAACCATCCGCTTTGGAGATTTTCCGCATG 19964
QY 33 HisGlu-----AsnArgAsnGlyThr 39
DB 19963 CACGACCGGTATTCGACCGACCCGCAATACGCGGCACTGGAATAGCTTACAGAAAGAACCA 19904
QY 40 TyrGlnTPPIleValLysProLysSerValAsnTyrGlnPheLysThrAsnThrHisVal 59
DB 19903 TCACCATGAAACAAAGAGTGTGCGGAAGCCGACGCGAAGCTCCGCGTCCG-ATY 19845
QY 60 ProLYLeuGlyValMetLeuValGlyTyrGlyGluAsnAsnGlySerThrLeuThrGly 79
DB 19844 CCGGCTTGGGCGCTGTC-----TCCACCACTGATGGCG 19809
QY 80 GlyValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAla 99
DB 19808 GGTGTGAGCTGGCGCGCAAGGCG-----AAGGCT 19779
QY 100 AsnTyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySer---PheGlnGly 118
DB 19778 CACCCCATTTGGGTGCTGCTACCGACAGATGGGACCGCCCGCTGGGAGACGAGCCGACGCG 19719
QY 119 GIUGluIleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleVal 138
DB 19718 CGCACCGTG-----AACCTCAACGAGCTGGTGTCTTGGCGGAGCTGAAGACCTTGCC 19665
QY 139 PheGlyGlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysVal 158
DB 19664 TTCCGCGCTGGACATCATCGTATGAGGACCGCATACGAAGTGGCGCGCTCGGCGCGT 19605
QY 159 PheAspIleAspLeuGluLysGlnLeuArgProTyrMetGluSerMetValProLeuPro 178
DB 19604 CTGACGCAACAGCACTGGAAGAGGTGAAGCGCTTCTCCAGAGCATCAAGCCGAAGAG 19545
QY 179 GlyIleTyrAspProAspPheIleAlaIleAsnGlnGluGluArgAlaAsnAsnAlaIle 198
DB 19544 GCGGTGACAGACCTGAGTGTGTGGCGCGCATCGAGGCC-----AACCAATC 19497
QY 199 LysGlyThrLys-----GlnGluGlnAlaGlnIleIleIleLysAspIleLysAlaPhe 216
DB 19496 AAGGCCACCAAGACGACACCGGAGAGATCGAAGCGCTGCCGACGATCCGCACTTC 19437
QY 217 LysGluAlaThrLysValAspLysValValValLeuTyrPheAlaAsnThrGluArgTyr 236
DB 19436 AAGAAAGAGCTCAACGCGACGCGCGGTATGAGGTGTCAGACGAGCTGACGACTTC 19377
QY 237 SerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArg 256
DB 19376 GGTCCGCTGCCCGAGTCTTCAAG--ACGCTGGCGGCTTCGAGAGAGCGCTGACAG 19320
QY 257 AsnGluAlaGluIleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnVal 276
DB 19319 AAGAGCCCGGACATCAACCCACGCGCTGTACACCTAGCGGCGCATCAAGAGGCGGTG 19260
QY 277 ProPheIleAsnGlySerProGln--AsnThrPheValProGlyLeuIleAspLeuAla 295
DB 19259 CCTTTCGCAAGCGCACCGCCCAAGCGCGTGAACGCGGCGCTCCAGAGATGGCC 19200
QY 296 IleAlaArgAsnThrIleIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLys 315
DB 19199 AAGCAGAGTGGTGGCGCGTGGCGCGCGCATCAAGACGCGGCAACATGATGAAG 19140
QY 316 SerValLeuValAspPheLeuValGlyValaGlyIleLysProThrSerIleValSerTyr 335
DB 19139 ACGGTATCGGCGCGCGCTTAAGCGCCGATGCTGGTCTGATGCTGCTTCCACAC 19080
QY 336 AsnHisLeuGlyYAAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLys 355
DB 19079 AACATCTCGGCAACCGGACGCGGAGGTGCTGACGACCGCGGCGCTTCAAGGCCAAG 19020
QY 356 GluIleSerLysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlu 375

```



```
Db 19019 GAAGTACCAAGTGGAGTGTGACACATCTG-----CAG 18981
Qy 376 ProGlyGlnHisProAsp-----HisValValIleLeuSerThrAla 389
Db 18980 CCGGACCTGTACCGGACCTGTACAGAGTACGGGACAGAGTGGCCATCCACTACTAC 18921
Qy 390 ProTyrValIleGlyAspSerLeuArgAlaMetAspGluTyrThrSerGluIlePheMetGly 409
Db 18920 CCGCCCCGGGGGACGGCAAGAGAGGTTGGGACCAACATCGACATCCGAGTGGCTGGC 18861
Qy 410 GlyLysAsnThrIleValIleuHisAsnThrCysGluAspSerLeuLeuAlaProIle 429
Db 18860 TATCCGATGACGATCAGAGTCAACTCTCTGCGCCGACATCCATCTGCGCGCTG 18801
Qy 430 IleLeuAspLeuValIleuLeuAlaGluLeuSerThrArgGlnPheLysAlaGluAsn 449
Db 18800 GTGCTGACATCGCGCTGTCTGACCTGGCCAGCGGCTGGAGTGGCGGCATCCAG 18741
Qy 450 GluGlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyrLeuThrLysAla 469
Db 18740 GAGTGG-----ATGCTCTTACTTCAAGACC 18714
Qy 470 ProLeuValIleProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeu 489
Db 18713 CCCATGGCGGACGCGGCTGTGCGGGTGGAGCAGACCTTTCATCCAGTCCAGCAAGCTG 18654
Qy 490 GluAsnIleMetArgAlaCysValIleuAlaProGluAsnAsnMetIleGluGluTyr 509
Db 18653 AAGAACACGCTGCGGCTGTCGCGGGCGAGAGCCCATCACCACCTGGGGCTGCATTAT 18594

RESULT 14
US-09-902-540-5256 Application US/09902540
; Sequence 5256, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Steven J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5256
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5256

Alignment Scores:
Pred. No.: 4,936-56 Length: 1335
Score: 528.00 Matches: 135
Percent Similarity: 49.24% Conservative: 91
Best Local Similarity: 29.41% Mismatches: 187
Query Match: 20.07% Indels: 46
Db: 4 Gaps: 11

US-10-718-952-12 (1-510) x US-09-902-540-5256 (1-1335)
Qy 61 LysLeuGlnValIleMetLeuValIleTyrGlyAlaAsnGlySerThrLeuThrGlyGly 80
Db 37 AAGCTCGCGGCTCTATTCCGGGCTCGGCGCTGTGTC--ACCACCTGATGGCGGGT 93
Qy 81 ValIleAlaAsnArgGluAspIleSerThrPalaThrLysAspLysIleGlnGlnAlaAsn 100
Db 94 GTCGAGCTCGCGCGGCAAGGC-----AAGGTCAAC 123
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySer--PheGlnGlyGlu 119
|||||
|||||
|||||
```

```
Db 124 CCGATTGGGTCGCTCACGACAGATGGGACCGCGCGCTGGGGAAAGCGGACCGGCGGC 183
Qy 120 GluIleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAlaIleValPhe 139
Db 184 ACCGTG-----AAGCTCAACGAGCTGTGCTTGTGGCGGAGCTGAAGAGCTTGCTTC 237
Qy 140 GlyIleTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPhe 159
Db 238 GCGCCCTGGGACATCATCCGCGAGGACGCATACGAAAGTGGCGCTGCGCGGCTGCTC 297
Qy 160 AspIleAspLeuGlnLysGlnLeuArgProTyrMetGluLeuMetValProLeuProGly 179
Db 298 AGCCACAAAGCACTGGAGAGGTAAAGCCCTTCTCCAGACATCAAGCCGACAGGCG 357
Qy 180 IleTyrAspProAspPheIleAlaIleAsnGlnGluGluArgAlaAsnValIleLys 199
Db 358 GTGACACACCTGACTTGTGCGCGCGCATGAGGC-----AACCATCAAG 405
Qy 200 GlyThrLys-----GlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLys 217
Db 406 GCCACAAAGACGACCGGAGAGCATGAAAGCGCTGGCGCGGACATCCGCACTTCAAG 465
Qy 218 GluAlaThrLysValAspLysValValIleLeuThrPheAlaAsnThrGluArgTyrSer 237
Db 466 AAGAGGCTCAACGGACGCGCGCTGATGTGTGACAGCAGCGTGGAGACCTTCCGT 525
Qy 238 AsnLeuValIleGlyLeuAsnAspThrMetGluAsnLeuAlaIleValAspArgAsn 257
Db 526 CCGCTGCCCGAGTCTTCAAG--ACGCTGCGCGCTTCGAAAGCGCTGGAGAGAAC 582
Qy 258 GluAlaGluIleSerProSerThrLeuTyrAlaIleAcylValMetGluAsnValPro 277
Db 583 AGCCCGGACATCAACCCACCGCGCTGTACCTACGAGCGGCATCAAGAGAGGCGTCC 642
Qy 278 PheIleAsnGlySerProGln--AsnThrPheValProGlyLeuIleAspLeuAlaIle 296
Db 643 TTCCTGAAAGCGGACCGCCCAAGCGGACGTGACACGCGGCGCTCAGAGATGACCAAG 702
Qy 297 AlaArgAsnThrLeuIleGlyIleAspAspPheLysSerGlyGlnThrLysMetLysSer 316
Db 703 CAGAGATCGGTGGCGGCTGCGCGCGGACCTCAAGAGCGGCGGACAGATGATGAAGCG 762
Qy 317 ValLeuValAspPheLeuValIleGlyIleLysProThrSerIleValSerTyrAsn 336
Db 763 GTCATCGCGCGCGCTCAAGCGCGGACGTGGTGTGATGCTGATGCTTCCACCAAC 822
Qy 337 HisLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlu 356
Db 823 ATCTGGAACCGGACGCGGAGGTGCTGACACCGCGGCGCTTCAAGGCAAGGAA 882
Qy 357 IleSerLysSerAsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrGluPro 376
Db 883 GTCAACAAAGTGAAGCTGTGGAACCATCTG-----CAGCC 921
Qy 377 GlyLysHisProAsp-----HisValValIleLysTyrValPro 390
Db 922 GACCTGTACCGGACGTGTACAGAGTACGCGACAGAGTGGCGCATCTACCTACCG 981
Qy 391 TyrValIleAspSerLysArgAlaMetAspGluTyrThrSerGluIlePheMetGlyGly 410
Db 982 CCGCGCGCGGACGGGAGAGGAGGTTGGGACAAATCGACATCCAGGATGGCTGGCAT 1041
Qy 411 LysAsnThrIleValIleuHisAsnThrCysGluAspSerLeuLeuAlaProIleIle 430
Db 1042 CCGATGACAGATCAAGGTCAACTTCTCTGCGGACATCTCACTTCGCGCGCGCTGGTG 1101
Qy 431 LeuAspLeuValIleuLeuAlaGluLeuSerThrArgIleGlnPheLysAlaGluAsnGlu 450
Db 1102 CTGACATCGCGCTTCTGTGACCTGGACCAAGCGCTGAGTGGCGGCGCATCAGAG 1161
Qy 451 GlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyrLeuThrLysAlaPro 470
Db 1162 TGG-----ATGCTCTTCTACTTCAAGACCC 1188
|||||
|||||
|||||
```

QY 471 LeuValProGlyThrProValAlaSerLysGluArgAlaMetLeuGlu 490
DB 1189 ATGGCCGACCGGCTGCGGAGACGACCTTTCATCCAGCTCAACAGCTGAAG 1248
QY 491 AsnIleMetArgAlaCysValGlyLeuAlaProGluAsnAsnMetIleLeuGluTyr 509
DB 1249 AACACGCTGCGGCTGCGGAGACCCCATCACCACCTCGGCTGCACTAT 1305

RESULT 15
US-09-313-294A-4684
; Sequence 4684, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4684
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348858H1
; NAME/KEY: unsure
; LOCATION: 185, 272, 275, 281, 283, 288
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4684

Alignment Scores:
Pred. No.: 9,51e-11 Length: 294
Score: 160.00 Matches: 28
Percent Similarity: 82.61% Conservative: 10
Best Local Similarity: 60.87% Mismatches: 8
Query Match: 6.08% Indels: 0
DB: 4 Gaps: 0

US-10-718-952-12 (1-510) x US-09-313-294A-4684 (1-294)

QY 1 MetPheIleGluAsnPhelyValGluSerProAsnValLysTyrThrGluThrGluIle 20
DB 125 ATGTCATCGAGAGCTTCGCGTCCGAGAGCCCCACGTCGGTACGCGCCGACGAGATC 184
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 185 NTCTCGAGGTACCGGTACGACGAGGAGCTGTGTGCACAGAGCCAGAGACGCGCTCC 244
QY 41 GlnTPrIleValLysPro 46
DB 245 CGCTGGCTGTCGCGCCC 262

Search completed: June 7, 2005, 22:59:18
Job time : 307 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:49:11 ; Search time 7680.29 Seconds
(without alignments)
11103.899 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttctctctcttctgt.....aatgtaataatcttctgt 1760

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1760	100.0	1782	6	BD075266 Soybean p
2	1777	97.6	1739	8	AY038802 Glycine m
3	1482	84.2	1533	6	BD075269 Soybean p
4	1454	82.6	1791	8	AF293970 Glycine m
5	65	3.7	62713	8	AP006418 Lotus cor
6	59	3.4	2994	8	AF282263 Phaseolus
7	45	2.6	1391	8	AY005128 Actinidia
8	40	2.3	119858	8	AP006419 Lotus cor
9	35	2.0	1978	8	CPINOIG C.paradisi
10	32	1.8	2053	8	UC125351 Mesembryant
11	32	1.8	143012	8	AC125389 Medicago
12	26	1.5	1533	8	AY33824 Xerophyta
13	26	1.5	1564	8	AY096554 Arabidops
14	26	1.5	1781	6	AX054630 Sequence
15	26	1.5	1781	8	BN06307 Brasica na
16	26	1.5	1845	8	AY028259 Avicennia
17	26	1.5	1863	8	AY065415 Arabidops
18	26	1.5	1986	8	AF433879 Suaeda ma
19	26	1.5	80088	8	AT131P16 Arabidops

20	24	1.4	35	6	BD075267 Soybean p
21	21	1.4	149633	2	AC087428 Homo sapi
22	24	1.4	187003	9	AC023353 Homo sapi
23	23	1.3	1535	6	BD073472 Regulated
24	23	1.3	1538	8	AF120146 Trifolium
25	23	1.3	1538	8	AF120147 Trifolium
26	23	1.3	1538	8	AF120148 Trifolium
27	23	1.3	1539	8	AF120148 Trifolium
28	23	1.3	1578	6	AF122340 Porteresia
29	23	1.3	1578	6	E13612 gDNA encodi
30	23	1.3	1578	6	E17392 gDNA encodi
31	23	1.3	1665	6	BD073470 Regulated
32	23	1.3	1665	8	AF056326 Zea mays
33	23	1.3	1868	8	AB012107 Oryza sat
34	23	1.3	1914	8	AK103501 Oryza sat
35	23	1.3	1915	8	AK058750 Oryza sat
36	23	1.3	1931	6	AK137808 Sequence
37	23	1.3	1931	6	AR170167 Sequence
38	23	1.3	1931	6	AB032073 Nicotiana
39	23	1.3	1936	8	AB059557 Avena sat
40	23	1.3	1950	6	E27176 Novel INPS
41	23	1.3	1959	8	AB009881 Nicotiana
42	23	1.3	1959	8	BT013505 Lycopersi
43	23	1.3	3546	6	AR137811 Sequence
44	23	1.3	3546	6	AR137812 Sequence
45	23	1.3	3546	6	AR170170 Sequence
	23	1.3	3546	6	AR170171 Sequence

ALIGNMENTS

RESULT 1	BD075266	1782 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD075266	Soybean plant producing seeds with reduced levels of raffinose			
DEFINITION	BD075266	saccharides and phytic acid.			
ACCESSION	BD075266.1	GI:22620869			
VERSION	BD075266	JP 2001519665-A/1.			
KEYWORDS	JP 2001519665-A/1.	unidentified			
SOURCE	unidentified	unclassified			
ORGANISM	unclassified.	unclassified.			
REFERENCE	1 (bases 1 to 1782)				
AUTHORS	Hitz,W.D. and Sebastian,S.A.				
TITLE	Soybean plant producing seeds with reduced levels of raffinose				
JOURNAL	Patent: JP 2001519665-A 1 23-OCT-2001;				
COMMENT	RI DU PONT DE NEMOURS AND CO				
	Soybean line IRI3				
	PN JP 2001519665-A/1				
	PD 23-OCT-2001				
	PF 07-APR-1998 JP 1998543012				
	PR 08-APR-1997 US 08/835751				
	PI WILLIAM DEAN HITZ,SCOTT ANTHONY SEBASTIAN				
	PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC				
	Strandedness: Double;				
	CC Topology: Linear;				
	CC Soybean plant producing seeds with reduced				
	levels of raffinose				
	CC and phytic acid				
	FM Key Location/Qualifiers				
	FT CDS 54..1586.				
FEATURES	source	Location/Qualifiers			
	1..1782	Location/Qualifiers			
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN	Query Match	100.0%; Score 1760; DB 6; Length 1782;			
	Best Local Similarity	100.0%; Pred.No. 0;			
	Matches 1760; Conservative	0; Mismatches	0; Indels	0; Gaps	0;

```

OY      1 CTCTCTTATTCCTTTTGTATTTCAATTCATTCTTAATCTTGTGAATAAATATGTTCA 60
DB      1 CTCTCTTATTCCTTTTGTATTTCAATTCATTCTTAATCTTGTGAATAAATATGTTCA 60
OY      61 TCGAATATTTAAGGTGAGTGTCTTAATGTGAAGTACACGAGACTGAGATTGAGTCCG 120
DB      61 TCGAATATTTAAGGTGAGTGTCTTAATGTGAAGTACACGAGACTGAGATTGAGTCCG 120
OY      121 TGTACACTAGCAAAACACCGAATCTTGTACAGAGAACGGAATGGCACTATCAGTGA 180
DB      121 TGTACACTAGCAAAACACCGAATCTTGTGTACAGAGAACGGAATGGCACTATCAGTGA 180
OY      181 TTGTGAACCCAAATCTGTCAATTCGAATTTTAAACCAATCCATGTTCTTAATTAG 240
DB      181 TTGTGAACCCAAATCTGTCAATTCGAATTTTAAACCAATTCATGTTCTTAATTAG 240
OY      241 GGGTAATGCTTGTGGGTGGGGTGGAAACAAGGCTCAACCTCAACGGTGGTGTATTG 300
DB      241 GGGTAATGCTTGTGGGTGGGGTGGAAACAAGGCTCAACCTCAACGGTGGTGTATTG 300
OY      301 CTAAACGAGAGGACATTTATGAGGCTCAAAAGACAAATTCACAAAGCAATTAATTG 360
DB      301 CTAAACGAGAGGACATTTATGAGGCTCAAAAGACAAATTCACAAAGCAATTAATTG 360
OY      361 GGTCCCTCAACCAAGCTGAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAAATCTATG 420
DB      361 GGTCCCTCAACCAAGCTGAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAAATCTATG 420
OY      421 CCCCATTCAGAGACCTGCTTCCATGTGTTAACCTGACGACATTTGTGTGGGGATGG 480
DB      421 CCCCATTCAGAGACCTGCTTCCATGTGTTAACCTGACGACATTTGTGTGGGGATGG 480
OY      481 AATATGACAAATGAACTGTGCTGATGCAATGAGGCAAGGCTTGTGATGATGAT 540
DB      481 AATATGACAAATGAACTGTGCTGATGCAATGAGGCAAGGCTTGTGATGATGAT 540
OY      541 TGCAGAGAGTGTGAGGCTTACATGGAATCCATGCTTCACTCCCGGAACTATAGAC 600
DB      541 TGCAGAGAGTGTGAGGCTTACATGGAATCCATGCTTCACTCCCGGAACTATAGAC 600
OY      601 CGGATTTCAATGCTGCCAACCAAGAGAGCTGCCAACACGTCATCAAGGGACAAAGC 660
DB      601 CGGATTTCAATGCTGCCAACCAAGAGAGCTGCCAACACGTCATCAAGGGACAAAGC 660
OY      661 AAGAGCAAGTTCAACAAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACCAAGTGG 720
DB      661 AAGAGCAAGTTCAACAAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACCAAGTGG 720
OY      721 ACAAGGTGTTGACTGTGGAATGCGCAACAGAGAGTACAGTATTTGTTGTGGGC 780
DB      721 ACAAGGTGTTGACTGTGGAATGCGCAACAGAGAGTACAGTATTTGTTGTGGGC 780
OY      781 TTAATGACACCAATGAGAAATCTTTGGCTGCTGTGACAAATGAAGCTGAGATTTCTC 840
DB      781 TTAATGACACCAATGAGAAATCTTTGGCTGCTGTGACAAATGAAGCTGAGATTTCTC 840
OY      841 CTTCACCTTGTATGTCATTTGCTTGTGTATGAAATGTTCTTCAATTAATGAGGCC 900
DB      841 CTTCACCTTGTATGTCATTTGCTTGTGTATGAAATGTTCTTCAATTAATGAGGCC 900
OY      901 CTCAGAAACCTTTTACAGAGGCTGATTGATCTTCCATCGGAGAAACATTTGATTG 960
DB      901 CTCAGAAACCTTTTACAGAGGCTGATTGATCTTCCATCGGAGAAACATTTGATTG 960
OY      961 GTGAGAGTACTTCAAGAGTGTCAAGCAAAATGAATCTGTGTGTGATTCTTCTTG 1020
DB      961 GTGAGAGTACTTCAAGAGTGTCAAGCAAAATGAATCTGTGTGTGATTCTTCTTG 1020
OY      1021 TGGGGGCTGGTATCAAGCAACATCTATATGCTCAATCAACATCTGGGAAACATGATG 1080
DB      1021 TGGGGGCTGGTATCAAGCAACATCTATATGCTCAATCAACATCTGGGAAACATGATG 1080

```

```

OY      1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGAAATCTCCAAAGACAACTGTG 1140
DB      1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGAAATCTCCAAAGACAACTGTG 1140
OY      1141 TTGATGATATGTTCAACAGCAATGCGATCTCTATGAGCTGTGTGAACATCCAGCCATG 1200
DB      1141 TTGATGATATGTTCAACAGCAATGCGATCTCTATGAGCTGTGTGAACATCCAGCCATG 1200
OY      1201 TTGTTGTTATTAAGATATGCTTACATGAGGGGACAGCAAGAGACCATGATGATGACA 1260
DB      1201 TTGTTGTTATTAAGATATGCTTACATGAGGGGACAGCAAGAGACCATGATGATGACA 1260
OY      1261 CTTCAGAGATATTCATGAGGTGAAAGACACCATTTGTTTGCAACAATGCGAGATT 1320
DB      1261 CTTCAGAGATATTCATGAGGTGAAAGACACCATTTGTTTGCAACAATGCGAGATT 1320
OY      1321 CCCTCTTAGCTGCTCTTATTTATCTTGACATTTGCTCTTCTTGCTAGACTCAGACTTAAGA 1380
DB      1321 CCCTCTTAGCTGCTCTTATTTATCTTGACATTTGCTCTTCTTGCTAGACTCAGACTTAAGA 1380
OY      1381 TCGAGTTTAAAGCTGAAATGAGGAAATTCACATTCATCCAGCCAGTGGTACATCC 1440
DB      1381 TCGAGTTTAAAGCTGAAATGAGGAAATTCACATTCATCCAGCCAGTGGTACATCC 1440
OY      1441 TCAGTACCTCAACCAAGGCTCCTGTGTTCACCCGAGTACACAGTGTGAATGCAATTGT 1500
DB      1441 TCAGTACCTCAACCAAGGCTCCTGTGTTCACCCGAGTACACAGTGTGAATGCAATTGT 1500
OY      1501 CAAACAGGCTCAATGCTGTGAAAAATATATGAGGGCTTTGTGTGATTTGGCCCAAGAGA 1560
DB      1501 CAAACAGGCTCAATGCTGTGAAAAATATATGAGGGCTTTGTGTGATTTGGCCCAAGAGA 1560
OY      1561 AATAATGATTTCTGAGTACAAAGTGAAGAGTGGGACCGAAGAAATATATGTTGGGGTGA 1620
DB      1561 AATAATGATTTCTGAGTACAAAGTGAAGAGTGGGACCGAAGAAATATATGTTGGGGTGA 1620
OY      1621 CCTAGCTGAATGTTTATGTTATATATATGTTGCTTATTAATTTTGCAGTGAATTGAA 1680
DB      1621 CCTAGCTGAATGTTTATGTTATATATATGTTGCTTATTAATTTTGCAGTGAATTGAA 1680
OY      1681 TGCATCAGCTTCATTAATGCTTTAGAGCGGGGCAATATCTGTTTACTGGAACATGAATG 1740
DB      1681 TGCATCAGCTTCATTAATGCTTTAGAGCGGGGCAATATCTGTTTACTGGAACATGAATG 1740
OY      1741 AATGTAGTATTAATTTTGTGT 1760
DB      1741 AATGTAGTATTAATTTTGTGT 1760

```

RESULT 2
 AY038802 1739 bp mRNA linear PLN 16-JUL-2001
 LOCUS
 DEFINITION Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.
 ACCESSION AY038802
 VERSION AY038802.1 GI:14764465
 KEYWORDS
 SOURCE
 ORGANISM
 Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 1739)
 REFERENCE
 AUTHORS Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.
 TITLE Biochemical and Molecular Characterization of a Mutation that
 Confers a Decreased Raffinose and Phytic Acid Phenotype on
 Soybean Seeds
 JOURNAL Unpublished
 AUTHORS Carlson,T.J. and Hitz,W.D.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402,
 Wilmington, DE 19880-0402, USA

FEATURES
source

CDS

```

Location/Qualifiers
1..1739
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="M9e"
/db_xref="taxon:3847"
/dev_stage="developing embryo"
11..1543
/EC_number="5.5.1.4"
/function="cyclizes glucose 6-phosphate to
L-myo-inositol-1-phosphate"
/note="mIPS-1A"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/db_xref="GI:1476446"
/translation="MPIENPKVECPNVKYTEIEISVYVETTELVENHNGTYOMIV
KKSIVYERKTNHVKLGLVLMVGWNGSITLGVIANREGISWATOKIOANVF
GSLTOSAIRVSGSFOGEEIYAPFKSLPVPNPDIIVFGMDISNMILADAMRAKVPD
IDLOKOLRYMESMLPLRGIYDPDPIAANDEERANVIGTQEOVOQIHKIKAFKE
ARKVDKVVLTANTERYNSLVGLNDTMENTLAVDREAEISSTLYAIACMENV
PFINSPQNTFVPLIDLAIANRTLIGDDFSGQTKMSVLDFLVGAGIPTSIVS
VYHLGNNDGMLSAPOFRSKISKSNVDDMVNSNAIYEBGEHDHYVVIKYVYV
GDSKRAMDEYTSIEIFMGKSTIVLHNTCDSLLAEPILDLVLAEISRIEFKANE
GKHSFHPVATILISVLTAKAIVPDPPTPVNALSQKAMLENTIRACVGLAPENMILE
YK"

```

ORIGIN

Query Match 97.6%; Score 1717; DB 8; Length 1739;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

44 GTGAAAAATATGTTTCATCGAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCA 103
1 GTGAAAAATATGTTTCATCGAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCA 60
104 GACTGAGATTCAGTCCGTGTAACAACGAAACCAACCACTTGTCAGAGAACAGGAA 163
61 GACTGAGATTCAGTCCGTGTAACAACGAAACCAACCACTTGTCAGAGAACAGGAA 120
164 TGGCACTTACATGATGATTTGTCAAAACCAATCTGTCAATAAGATTTTAAACCAACAT 223
121 TGGCACTTACATGATGATTTGTCAAAACCAATCTGTCAATAAGATTTTAAACCAACAT 180
224 CCATGTTCTTAATTAGGGGTAATGCTTGTGGGTTGGGGTGAACAACGGCTCAACCT 283
181 CATGTTCTTAATTAGGGGTAATGCTTGTGGGTTGGGGTGAACAACGGCTCAACCT 240
284 CACCGGTGATGTTATGCTTAAACGAGAGGCAATTTGATGGGCTAACAAAGACAAAGTTCA 343
241 CACCGGTGATGTTATGCTTAAACGAGAGGCAATTTGATGGGCTAACAAAGACAAAGTTCA 300
344 ACAAGCAATTAATTGGGCTCCCTCAACCAAGGCTCAGATTCAGAGTTGGGCTCTTCA 403
301 ACAAGCAATTAATTGGGCTCCCTCAACCAAGGCTCAGATTCAGAGTTGGGCTCTTCA 360
404 GAGAGAGAAATCTATGCCCCATTTCAAGAGCTGCTTCAATGTTAACTTACCGACAT 463
361 GAGAGAGAAATCTATGCCCCATTTCAAGAGCTGCTTCAATGTTAACTTACCGACAT 420
464 TGTGTTTGGGGATGAGGATATCGCAACATGAACTTGGCTGATGCAATGGCCAGGGCAAA 523
421 TGTGTTTGGGGATGAGGATATCGCAACATGAACTTGGCTGATGCAATGGCCAGGGCAAA 480
524 GGTGTTTGCATGATTTTGCAGAGAGATGAGGCTTACATGGAATCCATGCTTCACT 583
481 GGTGTTTGCATGATTTTGCAGAGAGATGAGGCTTACATGGAATCCATGCTTCACT 540
584 CCCCGGAATCTATGATCCGATTTCAATGCTGCAACCAAGAGAGCGTCCAAACGCT 643
541 CCCCGGAATCTATGATCCGATTTCAATGCTGCAACCAAGAGAGCGTCCAAACGCT 600
644 CATCAAGGCAACAAGAGAGAGTTCACAAATCATCAAGAGATCAAGGCGTTTAA 703

```

```

DB 601 CATCAAGGCAACAAGAGAGAGTTCACAAATCATCAAGAGATCAAGGCGTTTAA 660
704 GGAAGCCACCAAGATGAGACAGAGTGTGTTACTGTGATCTGCCACACAGAGAGTACAG 763
661 GGAAGCCACCAAGATGAGACAGAGTGTGTTACTGTGATCTGCCACACAGAGAGTACAG 720
764 TAATTTGTTGTTGGGCTTAAATGACACCAATGAGAAATCTCTGGGCTGCTGAGACGAAA 823
721 TAATTTGTTGTTGGGCTTAAATGACACCAATGAGAAATCTCTGGGCTGCTGAGACGAAA 780
824 TGAAGCTGATTTCTCTCTTCCACCTTGATGACCAATGCTTGTATGAGAAATGTTCC 883
781 TGAAGCTGATTTCTCTCTTCCACCTTGATGACCAATGCTTGTATGAGAAATGTTCC 840
884 TTTTCATTAATGGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTCCATGCC 943
841 TTTTCATTAATGGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTCCATGCC 900
944 GAGGAACACTTTGATTTGTGTGAGATGACCTTCAAGAGTGTGACCAAAATGAAATCTGT 1003
901 GAGGAACACTTTGATTTGTGTGAGATGACCTTCAAGAGTGTGACCAAAATGAAATCTGT 960
1004 GTTGGTTGATTTCTTGTGGGGCTGTGATCAAGCCAAATCTATGATGATCAACCA 1063
961 GTTGGTTGATTTCTTGTGGGGCTGTGATCAAGCCAAATCTATGATGATCAACCA 1020
1064 TCTGGGAAAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1123
1021 TCTGGGAAAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1124 CTCGAAGAGCAACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1183
1081 CTCGAAGAGCAACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1184 TGAACATCCAGACCAATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1243
1141 TGAACATCCAGACCAATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1244 AGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1303
1201 AGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1304 CAACACATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1363
1261 CAACACATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1364 TGAAGCTGAGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1423
1321 TGAAGCTGAGACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1424 CCCAGTTGCTACATCTCTCAGTACCTTCAACAGGCTCCTCTGTGTTCAACCGGGTACAC 1483
1381 CCCAGTTGCTACATCTCTCAGTACCTTCAACAGGCTCCTCTGTGTTCAACCGGGTACAC 1440
1484 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1543
1441 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1544 TGAATTTGGCCCAAGAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1603
1501 TGAATTTGGCCCAAGAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1604 TAATATAGTTGGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1663
1561 TAATATAGTTGGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1664 TTGCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1723
1621 TTGCAAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
1724 TACTAGGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1760

```

DB 1681 TACTAGAACATGAATGAATGATATTAATTTGTGT 1717

RESULT 3
BD075269

LOCUS BD075269 1533 bp DNA linear PAT 27-AUG-2002

DEFINITION Soybean plant producing seeds with reduced levels of raffinose saccharides and phytic acid.

ACCESSION BD075269

VERSION BD075269.1 GI:22620872

KEYWORDS JP 2001519665-A/4.

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 1533)
Hitz, W.D. and Sebastian, S.A.
Soybean plant producing seeds with reduced levels of raffinose saccharides and phytic acid
Patent: JP 2001519665-A 4 23-OCT-2001;

JOURNAL EI DU PONT DE NEMOURS AND CO

COMMENT OS Soybean line LR33
PN JP 2001519665-A/4
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC
Strandness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced levels of raffinose
CC saccharides
FH key Location/Qualifiers
FT CDS 1..1533.
Location/Qualifiers
1..1533
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 84.2%; Score 1482; DB 6; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 ATGTTATCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 113
DB 1 ATGTTATCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 60

QY 114 CAGTCGCTGAACACTACGAAACACCGAAGTGTTCAGAGAAAGAGAAATGGCACTAT 173
DB 61 CAGTCGCTGAACACTACGAAACACCGAAGTGTTCAGAGAAAGAGAAATGGCACTAT 120

QY 174 CAGTCGATTTGTCAAACTCAATCTGTCAATAGAAATTTTAAACCAACATCCATGTTCT 233
DB 121 CAGTCGATTTGTCAAACTCAATCTGTCAATAGAAATTTTAAACCAACATCCATGTTCT 180

QY 234 AAATTTGGGGTAATGCTTTGGGGTGGGGTGGAAACAAGGCTCAACCTCACCGGTGGT 293
DB 181 AAATTTGGGGTAATGCTTTGGGGTGGGGTGGAAACAAGGCTCAACCTCACCGGTGGT 240

QY 294 GTTATTGCTAACCGAAGGCAATTTATGGGCTACAAAGGACAAAGATTCAACAGCAAT 353
DB 241 GTTATTGCTAACCGAAGGCAATTTATGGGCTACAAAGGACAAAGATTCAACAGCAAT 300

QY 354 TACTTTGGCTCCCTCAACCGCTCAGCTATCCGATTGGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCGCTCAGCTATCCGATTGGGTCTTCCAGGAGAGAA 360

QY 414 ATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTTAACTTGAAGCAATGTTGGG 473
DB 361 ATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTTAACTTGAAGCAATGTTGGG 420

QY 474 GGATGGATATCAGCAATGAACTGTGATGCCATGGCCAGGGCAAGGTGTTGAC 533
DB 421 GGATGGATATCAGCAATGAACTGTGATGCCATGGCCAGGGCAAGGTGTTGAC 480

QY 534 ATCGATTTGCAAGACAGTTGAGGCTTTACATGAAATCATAGTCTTCCACTCCCGGAATC 593
DB 481 ATCGATTTGCAAGACAGTTGAGGCTTTACATGAAATCATAGTCTTCCACTCCCGGAATC 540

QY 594 TATGACCGGGAATTTATGTTGCTGCCAACCAAGAGAGCGTGCACAAAGCTCATCAAGGAC 653
DB 541 TATGACCGGGAATTTATGTTGCTGCCAACCAAGAGAGCGTGCACAAAGCTCATCAAGGAC 600

QY 654 ACAAGCAAGACAAAGTTCAACAAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 713
DB 601 ACAAGCAAGACAAAGTTCAACAAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 660

QY 714 AAAGTGACAAAGGTGTTGTAAGTGTGACCTGCCAACACAGAGAGTACATGTAATTTGGTT 773
DB 661 AAAGTGACAAAGGTGTTGTAAGTGTGACCTGCCAACACAGAGAGTACATGTAATTTGGTT 720

QY 774 GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGTGAGACAGAAATGAGGCTGAG 833
DB 721 GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGTGAGACAGAAATGAGGCTGAG 780

QY 834 ATTTCTCCTTCCACCTTGTATGCTCATTTGCTTGTATGAAATGTTCTTTCAATAT 893
DB 781 ATTTCTCCTTCCACCTTGTATGCTCATTTGCTTGTATGAAATGTTCTTTCAATAT 840

QY 894 GGAAGCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGGCATGCGGAGAACT 953
DB 841 GGAAGCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGGCATGCGGAGAACT 900

QY 954 TTGATTTGGAGATGACCTTCAAGAGTGTGTCAGACCAAAATGAATCTGTGTTGTTGAT 1013
DB 901 TTGATTTGGAGATGACCTTCAAGAGTGTGTCAGACCAAAATGAATCTGTGTTGTTGAT 960

QY 1014 TTCTTTGGGGGCTGTGATCAAGGCCAACATCTATAGTCAAGTTACAAACATCTGGGAAAC 1073
DB 961 TTCTTTGGGGGCTGTGATCAAGGCCAACATCTATAGTCAAGTTACAAACATCTGGGAAAC 1020

QY 1074 AATGATGTATGATATCTTTGGCTCCAAACTTTCCGTTCCAGAAATCTCCAAAGAC 1133
DB 1021 AATGATGTATGATATCTTTGGCTCCAAACTTTCCGTTCCAGAAATCTCCAAAGAC 1080

QY 1134 AACGTTGTATGATATGATGTCACAGCAATGCCATCTATAGGCTGTGTAACATCCA 1193
DB 1081 AACGTTGTATGATATGATGTCACAGCAATGCCATCTATAGGCTGTGTAACATCCA 1140

QY 1194 GACCATGTTGTTGATTAATGATGTGCTTACGTAGGGGACAGCAAGAGACCATGAT 1253
DB 1141 GACCATGTTGTTGATTAATGATGTGCTTACGTAGGGGACAGCAAGAGACCATGAT 1200

QY 1254 GAGTACACTTCAAGATATTTACAGGGTGAAGAGACACATTTGTTTGCACAAACATGTC 1313
DB 1201 GAGTACACTTCAAGATATTTACAGGGTGAAGAGACACATTTGTTTGCACAAACATGTC 1260

QY 1314 GAGGATTCCTCTTGAAGTCTCTTATTAATCTTGAAGCTTGGTCTTGTGTAAGCTCAGC 1373
DB 1261 GAGGATTCCTCTTGAAGTCTCTTATTAATCTTGAAGCTTGGTCTTGTGTAAGCTCAGC 1320

QY 1374 ACTAGATCGAGTTTAAAGCTGAAATGAGGAAATTTCACTATTCACACCAAGTTGT 1433
DB 1321 ACTAGATCGAGTTTAAAGCTGAAATGAGGAAATTTCACTATTCACACCAAGTTGT 1380

QY 1434 ACCATCTCAGCTACCTCAACAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAT 1493
DB 1381 ACCATCTCAGCTACCTCAACAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAT 1440

QY 1494 GCATTTGCAAGACAGCTGCAATGCTGGAATAACATTAAGAGGCTTGTGTTGATTTGGCC 1553
DB 1441 GCATTTGCAAGACAGCTGCAATGCTGGAATAACATTAAGAGGCTTGTGTTGATTTGGCC 1500

QY 1554 CCAGAGATACATGATTCGAGTACCAAGTGA 1586
DB 1501 CCAGAGATACATGATTCGAGTACCAAGTGA 1533

RESULT 4
AF293970
LOCUS AF293970
DEFINITION Glycine max myo-inositol-3-phosphate synthase (MIPS) mRNA, complete cds.
ACCESSION AF293970
VERSION AF293970.1 GI:13936690
KEYWORDS
SOURCE Glycine max (soybean)
ORGANISM Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 1791)
AUTHORS Hegeman, C.E., Good, L.L. and Grabau, E.A.
TITLE Expression of D-myo-inositol-3-phosphate synthase in soybean. Implications for phytic acid biosynthesis
JOURNAL Plant Physiol. 125 (4), 1941-1948 (2001)
MEDLINE 21196082
PUBMED 11299373
REFERENCE 2 (bases 1 to 1791)
AUTHORS Hegeman, C.E., Good, L.L. and Grabau, E.A.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA 24060, USA

FEATURES
source
1. 1791
/organism="Glycine max"
/mol_type="mRNA"
/culivar="Williams 82"
/db_xref="taxon:3847"
/dev_stage="immature cotyledons"
1. 1791
/gene="MIPS"
62. 1594
/gene="MIPS"
/note="isomerase"
/codon_start=1
/product="myo-inositol-3-phosphate synthase"
/protein_id="AAK49896.1"
/db_xref="GI:13936691"
/translation="MPEIENKRCVNRKTEETIGSVNYETTELVEHNRNGYOMIV
KPSVRYEFKTNIHVRLGVMLVGWGNNGSTLTGVIANREGISWATKDKIQQANVF
GSLTQASAIRVGSFOGEIYAPFKSLIPWNPDDIVFGGNDISNNMLADAMAKYFD
IDLOKQMRPYMESMLPLPGIYDPDLAANGEEANNVIKTKQEQOQIIKDKAKKE
ATKVDKVVVMTANTERYSNLVVGNDTMEMLAAVDRAEASIPSTLYAIACMVN
PFTNGSPONTFVPGILDIAARWTLIGGDFKSGOTKMSVIVPLVAGGIRKSTIVS
YHNLGNNDKMLSAFOTFRSKESKSNVDDMVNSNAILYBPEEHDDHVIVIKTYIV
GDSKRAMDEYTSIEFMWGKSTIVLHNTCESLAAPIILDVLALBELSTRIEKANE
GKHSPHPVATIIISYLTAKPLVPPTPGVNAISKORAMLENIRACVGLAPENMILE
YK"

ORIGIN
Query Match 82.6%; Score 1454; DB 8; Length 1791;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1754; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTCTCTTATTCCTTTGTAATTCATTCATTCCTTAATCTTTGTGAATAATATGTCA 60
DB 9 CTCTCTTATTCCTTTGTAATTCATTCATTCCTTAATCTTTGTGAATAATATGTCA 68

QY 61 TCGAGATTTTAAGTGTGAGTGTCTTAATGTGAAGTACACGAGACTGAGATTCAGTCCG 120
DB 69 TCGAGATTTTAAGTGTGAGTGTCTTAATGTGAAGTACACGAGACTGAGATTCAGTCCG 128

QY 121 TGTACAACTACGAAACCAACGAACTGTTCACAGAGAAAGAGAACTGCACTATCACTGGA 180

DB 129 TGTACAACTACGAAACCAACGAACTGTTCACAGAGAAAGAGAACTGCACTATCACTGGA 188
QY 181 TTGTCAACCCCAATTTGTCAAAATACGAATTTTAAACCAACATCCATGTTTCTTAATTAG 240
DB 189 TTGTCAACCCCAATTTGTCAAAATACGAATTTTAAACCAACATCCATGTTTCTTAATTAG 248

QY 241 GGGTAATGCTTGTGGGTGGGGTGGAAACAAGGCTCAACCTCAACGGTGGTGAATTG 300
DB 249 GGGTAATGCTTGTGGGTGGGGTGGAAACAAGGCTCAACCTCAACGGTGGTGAATTG 308

QY 301 CTAAACGAGAGGCAATTTTCATGAGGCTCAAAAGGACAAATTCACAAAGCAATTCCTTTG 360
DB 309 CTAAACGAGAGGCAATTTTCATGAGGCTCAAAAGGACAAATTCACAAAGCAATTCCTTTG 368

QY 361 GCTCCTCAACCAAGCTCAAGCTATCCGATGTTGGTCTTCCAGGAGAGAAATCTATG 420
DB 369 GCTCCTCAACCAAGCTCAAGCTATCCGATGTTGGTCTTCCAGGAGAGAAATCTATG 428

QY 421 CCCCATTCAGAGCCGCTTCCATGCTTAACCTGACGCAATGCTTGGGGATGGG 480
DB 429 CCCCATTCAGAGCCGCTTCCATGCTTAACCTGACGCAATGCTTGGGGATGGG 488

QY 481 ATATCAGCAACATGAACCTGCGTGAATCCATGAGCCAGGCAAGGCTTTGACATGATTT 540
DB 489 ATATCAGCAACATGAACCTGCGTGAATCCATGAGCCAGGCAAGGCTTTGACATGATTT 548

QY 541 TGCAGAGAGATTTAGAGCCCTTACATGAATTCATGCTTCACTCCCGGAATCTATGACC 600
DB 549 TGCAGAGAGATTTAGAGCCCTTACATGAATTCATGCTTCACTCCCGGAATCTATGACC 608

QY 601 CGGATTTCAATTGCTGCGCAACAAAGAGACGTGCCAAACGTCATCAAGGGGCAAAAGC 660
DB 609 CGGATTTCAATTGCTGCGCAACAAAGAGACGTGCCAAACGTCATCAAGGGGCAAAAGC 668

QY 661 AAGAGCAAGTTCAACAAATCATCAAAAGCATCAAAAGGCTTTTAAAGAAAGCCAAAGTGG 720
DB 669 AAGAGCAAGTTCAACAAATCATCAAAAGCATCAAAAGGCTTTTAAAGAAAGCCAAAGTGG 728

QY 721 ACAAGGTGTTGTAAGTGTGAGTGCACCAAGAGAGTACAGTAATTTGTTGTGGGCC 780
DB 729 ACAAGGTGTTGTAAGTGTGAGTGCACCAAGAGAGTACAGTAATTTGTTGTGGGCC 788

QY 781 TTATATGACACCATGAGAAATCTCTTGGCTGCTGTGACAGAAATGAGGCTGATTTCTC 840
DB 789 TTATATGACACCATGAGAAATCTCTTGGCTGCTGTGACAGAAATGAGGCTGATTTCTC 848

QY 841 CTTTCACTTGTATGCAATGCTGTGTTATGGAATAATGTCCTTCAATTAATGGAAGCC 900
DB 849 CTTTCACTTGTATGCAATGCTGTGTTATGGAATAATGTCCTTCAATTAATGGAAGCC 908

QY 901 CTCAGAACCTTTTGTACAGAGGCTGATGATCTTCCATCCGAGAGAACACTTTGATTTG 960
DB 909 CTCAGAACCTTTTGTACAGAGGCTGATGATCTTCCATCCGAGAGAACACTTTGATTTG 968

QY 961 GTGAGATGATTTCAAGAGTGTCAAGCAAAATGAATCTGTGTTGTTGATTTCTTTG 1020
DB 969 GTGAGATGATTTCAAGAGTGTCAAGCAAAATGAATCTGTGTTGTTGATTTCTTTG 1028

QY 1021 TGGGGCTGTATCAAGCAACATCTATGAGTGAACCAATCTGGGAAACATGATG 1080
DB 1029 TGGGGCTGTATCAAGCAACATCTATGAGTGAACCAATCTGGGAAACATGATG 1088

QY 1081 GTATGATCTTTTCGGCTCAACAACTTTCCGTTCCAGAGAAATCTCCAGAGCAACGTTG 1140
DB 1089 GTATGATCTTTTCGGCTCAACAACTTTCCGTTCCAGAGAAATCTCCAGAGCAACGTTG 1148

QY 1141 TTGATGATGATGATCAACAGCAATGCCATCTCTATGAGCTGTGTGAACATCAAGCATG 1200
DB 1149 TTGATGATGATGATCAACAGCAATGCCATCTCTATGAGCTGTGTGAACATCAAGCATG 1208

QY 1201 TTGTTCTTATTAATGATGCTTCTATGAGGAGACAGAGAGGCACTGATGATGA 1260

D	b		1209	TGTGTTATTATTAAGTAAVTGCTTCACGNAGGGGA CAGCAAGAAGGCATGATGATGACA	1268
O	y		1261	CTTCAGAGATTTTCATGGGTGGAAAAGCACCATTGTTTTCACACACATCGCAGCAT	1320
D	b		1269	CTTCAGAGATAATTCATGGGTGGAAAAGCACCATTTGTTTTCACACACATCGCAGCAT	1328
O	y		1321	CCCTCTTAGAGTGCTCTCTAATTCCTTGGACCTTGTCCTCTTCTGTGAGTCGACTAGAA	1380
D	b		1329	CCCTCTTAGAGTGCTCTCTAATTCCTTGGACCTTGTCCTCTTCTGTGAGTCGACTAGAA	1388
O	y		1381	TCGAGTTTAAAGCTGAAAAATGAGGAAAATTCACCTCATTTCCACCCAGTTGCTAACATCC	1440
D	b		1389	TCGAGTTTAAAGCTGAAAAATGAGGAAAATTCACCTCATTTCCACCCAGTTGCTAACATAC	1448
O	y		1441	TCACCTACCTCACCAGAGGCTCTGTGGTTCACCGGGGACACACAGTGTGTAATCATTTGT	1500
D	b		1449	TCACCTACCTCACCAGAGGCTCTGTGGTTCACCGGGGACACCGGGTGTGAATGATTTGT	1508
O	y		1501	CMAAGCAGCGGTGCAATGCTGGAAAAACATTAATGAGGGCTTGTTGTGTAATTTGCCCCAGAGA	1560
D	b		1509	CMAAGCAGCGGTGCAATGCTGGAAAAACATTAATGAGGGCTTGTTGTGTAATTTGCCCCAGAGA	1568
O	y		1561	ATTAACATGATTTCTGAGTACCAAGTGAAGCATGCGACCGAAGAAATTAATTAAGTTGGGGTAG	1620
D	b		1569	ATTAACATGATTTCTGAGTACCAAGTGAAGCATGCGACCGAAGAAATTAATTAAGTTGGGGTAG	1628
O	y		1621	CCTAGCTGAATGTTTTATGTTAATTAATTAATGTTTGTCTTAATTTTGCAGGTGAATGGA	1680
D	b		1629	CCTAGCTGAATGTTTTATGTTAATTAATTAATGTTTGTCTTAATTTTGCAGGTGAATGGA	1688
O	y		1681	TGCATCAGCTTCATTAATGCTTTAGAGCGGGGCAATATTCGTGTTTAAGTAAGCAATGAATG	1740
D	b		1689	TGCATCAGCTTCATTAATGCTTTAGAGCGGGGCAATATTCGTGTTTAAGTAAGCAATGAATG	1748
O	y		1741	AATGTAGTAAATTTTGTGT	1760
D	b		1749	AATGTAGTAAATTTTGTGT	1768
R	E	S	RESULT 5		
L	O	C	AP006418		
D	E	F	LOCUS		
I	N	I	DEFINITION	AP006418 62713 bp DNA linear PLN 22-JUL-2003	
I	N	I	LOCATION	Lotus corniculatus var. japonicus genomic DNA, chromosome 6,	
I	N	I	VERSION	clone:ljt23AO2, TM0306, complete sequence.	
I	N	I	KEYWORDS	AP006418	
I	N	I	SOURCE	AP006418.1 GI:31581049	
I	N	I	ORGANISM	HMG.	
I	N	I		Lotus corniculatus var. japonicus (lotus japonicus)	
I	N	I		Lotus corniculatus var. japonicus	
I	N	I		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
I	N	I		Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;	
I	N	I		rosids; eurosoids I; Fabales; Fabaceae; Papilionoideae; Loteae;	
I	N	I		Lotus.	
R	E	S	REFERENCE		
I	N	I	1	Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.	
I	N	I		Structural Analysis of a Lotus japonicus Genome. IV. Sequence	
I	N	I		Features and Mapping of seventy-three TAC clones which cover the	
I	N	I		7.5 Mb Regions of the Genome	
I	N	I		DNA Res. (2003) In press	
I	N	I		2 (bases 1 to 62713)	
I	N	I		Sato,S.	
I	N	I	JOURNAL	Direct Submission	
I	N	I	TITLE	Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,	
I	N	I		Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu,	
I	N	I		Chiba 292-0818, Japan (E-mail:ssato@kazusa.or.jp,	
I	N	I		URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),	
I	N	I		Fax:81-438-52-3934)	
I	N	I	FEATURES	Location/Qualifiers	
I	N	I	SOURCE	1..62713	

Query Match	Best local Similarity	100.0%	Pred. No. 1,2e-24	Matches 65	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1140	GTTGATGATATGGTCAACAGCAATGCCATCTCTATAGCGTGGTGAACATCCAGCCAT	1199					
Db	30144	GTTGATGATATGGTCAACAGCAATGCCATCTCTATAGCGTGGTGAACATCCAGCCAT	30203					
QY	1200	GTGTCT	1204					
Db	30204	GTGTCT	30208					
RESULT 6	AF2822263	2994 bp	DNA	linear	PLN 02-JUL-2001			
LOCUS	AF2822263	Phaeosolus vulgaris 1L-myo-inositol-1-phosphate synthase gene,						
DEFINITION	AF2822263	partial cds.						
ACCESSION	AF2822263							
VERSION	AF2822263.1	GI:14582466						
KEYWORDS								
SOURCE								
ORGANISM		Phaeosolus vulgaris						
		Phaeosolus vulgaris						
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;						
		rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaeoleae;						
		Phaeosolus.						
REFERENCE	1 (bases 1 to 2994)							
AUTHORS	Johnson,M.D., Lackey,K.H. and Pope,P.M.							
TITLE	Subcellular localization of 1L-myo-inositol-1-phosphate synthase							
JOURNAL	Unpublished							
REFERENCE	2 (bases 1 to 2994)							
AUTHORS	Pope,P.M. and Johnson,M.D.							
TITLE	Direct Submision							
JOURNAL	Submitted (24-JUN-2000) Biological Sciences, The University of							
	Alabama, P.O. Box 870344, Tuscaloosa, AL 35487, USA							
FEATURES		location/Qualifiers						
source	1..2994							
	/organism="Phaeosolus vulgaris"							
	/mol_type="genomic DNA"							
	/db_xref="taxon:3885"							
	join(<582..769,886..954,1047..1182,1335..1582,1662..1888							
	1965..2080,2243..2419,2503..2691,2795..>2860)							
	/product="1L-myo-inositol-1-phosphate synthase"							
	join(582..769,886..954,1047..1182,1335..1582,1662..1888,							
	1965..2080,2243..2419,2503..2691,2795..>2860)							
	/codon_start=1							
	/product="1L-myo-inositol-1-phosphate synthase"							
	/protein_id="AAK69514.1"							
	/db_xref="GI:14582467"							
	/translation="MFIEIMFKVESPNVKYSETIEGVINVEITELVHNENRNGAYOII							
	KPVSFYEKFTNTHVPLEKLVMLVGMGNGNSTLTGVIANRBEISVATQDKIQANVF							
	GSILQMSAIRVGSYGEIEYAPKSLIPWNPDIYVFGMDISNNMLADAMGAKVED							
	IDLQKRLPYMESWVPLPGIYDPDFIAANOESANNVIRGTTKEQVOVYIKDKERKA							
	ATAYDVRVVLWTRNTERYSNLVGLNDTBNILAADRGAELSPETLPAICVTENV							
	PTFINSYQNTFVPEGLIDFIAQKCLIGGDFKSGQTKMSVLVDPLVGGIKRPTSVS							
	YHNLGNNDGNLSAPQTFRSKEISKSNVVDVWNSVAILYEPEDHPDHVVVIRKYVY							
	AQSKRAMDEYTSIEIFMGGRNLTIVLHNTCEDSLIAPIIIDLVLALSTRIQFKAKE							
	GKHSHSPVAITLISVLTKAPLV"							

[illegible]

JOURNAL	DNA Res. (2003) In press
REFERENCE	2 (bases 1 to 119858)
AUTHORS	Sato, S.
TITLE	Direct Submision
JOURNAL	Submitted (07-MAY-2003) Shusui Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:satokazusa.or.jp, URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337), Fax:81-438-52-3934)
FEATURES	Location/Qualifiers
source	1. 119858
	/organism="Lotus corniculatus var. japonicus"
	/mol_type="genomic DNA"
	/variety="japonicus"
	/db_xref="taxon:34305"
	/chromosome="4"
	/clone="LJ143018"
	/clone_lib="LJT library"
	/note="TAC clone:TM0307-synonym: Lotus japonicus"
ORIGIN	
Query Match	2.3%; Score 40; DB 8; Length 119858;
Best Local Similarity	100.0%; Pred. No. 2e-10;
Matches	40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1175 TGAGCTGTGTGACATCCAGACCATGTTGTTATTAAAG 1214
Ddb	46349 TGAGCTGTGTGACATCCAGACCATGTTGTTATTAAAG 46310
RESULT 9	
CPINOIG	CPINOIG 1978 bp DNA linear Pln 26-JUN-1995
LOCUS	C.paradis1 (Macf) INO1 gene.
DEFINITION	Z32632
ACCESSION	Z32632.1 GI:602564
VERSION	INO1 gene.
KEYWORDS	Citrus x paradis1
SOURCE	Citrus x paradis1
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Sapindales; Rutaceae; Citrus.
	1 (bases 1 to 1978)
REFERENCE	Holland, D.
AUTHORS	Direct Submission
TITLE	Submitted (08-APR-1994) Doron Holland, Fruit Tree Breeding and Genetics, Volcani Research, Organisation Center, Bet Dagan, 50250, Israel
JOURNAL	
FEATURES	Location/Qualifiers
source	1. 1978
	/organism="Citrus x paradis1"
	/mol_type="genomic DNA"
	/db_xref="taxon:37656"
	/clone="C38"
	/rissue_type="fully expanded leaves"
	/dev_stage="juvenile plants"
	188..1711
	/gene="INO1"
	188..1711
	/gene="INO1"
	/codon_start=1
	/protein_id="CAA83565.1"
	/db_xref="GI:602565"
	/db_xref="GDB:P42802"
	/db_xref="UniProt/Swiss-Prot:P42802"
	/translation="MFLENKVESPVKTYDHSVYDYETTELVHENVNGTYOMIV KPTLVQEFKTDVHVKLGVLVAGMGNGNSTLTGGIARBEICVATKTDVQDANF KSTVQSAIRVSGYNGEELIYAPKSLIPWNPDIYFGWDIDMNLDAABAARVDF IDLOKOLPYMESNVLPGLGYDDPTLAANGSRANNVYKGTKEOMLOIKIDREFKGI KNNVDVYVLTANTERTERSNVYIGLVDTYSLASLDKNAELISPTLVAIACLENI PFINGSPONTFVGLIDLARRNCILGGDDFKSGQTKMSVLYVDFVGAQIKPSTIVS YNLGNDGNNLSAPQTFRSKEISKSNVVDWVSNVFPVGLVNTFRRWIKVYPVAI ERMADYTSIEFMWGSRTIVLHNTCEBSLLAAPLIIDLVLAIESTRIQKABEGFGR
gene	
CDS	

```

ORIGIN
      HSFHVPATILSYLTKAPLVPPGTVPVNAISKORAMLENILRACVGLADENNMLEYK"

Query Match      2.0%; Score 35; DB 8; Length 1978;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
246 ATGCTTGCTGGCTGGGTAACAACACGCGCTCAAC 280
|||||
380 ATGCTTGCTGGCTGGGTAACAACACGCGCTCAAC 414

RESULT 10
MCU32511      2053 bp mRNA linear PLN 11-JUL-1996
LOCUS
DEFINITION Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase
ACCESSION U32511
VERSION U32511
KEYWORDS U32511.1 GI:975887
SOURCE
ORGANISM Mesembryanthemum crystallinum (common iceplant)
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Alzooceae; Mesembryanthemum.
1 (bases 1 to 2053)
Ishlanti,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B.,
Jensen,R.G. and Bohnert,H.J.
Coordinate, transcriptional induction of myo-inositol metabolism
during environmental stress
Plant J. 9 (4), 537-548 (1996)
66208959
MEDLINE
PUBMED 8624516
REFERENCE
AUTHORS 2 (bases 1 to 2053)
Ishlanti,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B.,
Jensen,R.G. and Bohnert,H.J.
Direct Submision
Submitted (27-JUL-1995) Hans J. Bohnert, Biochemistry, University
of Arizona, Biosciences West 516, Tucson, AZ 85721, USA
FEATURES
source
1.2053
Location/Qualifiers
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
182..1720
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AB03687.1"
/db_xref="GI:975888"
/translacion="MFIESPKVESPNVKYTNESIVNYDTTELVENRKDAGGYOM
IKPKTVQAHFKDTRVPLKGLVLMVGNNGSTLGGVIANREGISNATKDKIOAN
YGSILQKASISIVSGFNGREIYAPFKSLIPNPNPDVVGMDISNMADAMTARF
FIDIDOKLRPFYHEHVPYPGIYDPFLIANGSRANVITKTKEQVERVYKIDIREF
KERNKVDVLMNTGNTERYSNVVGELNMTENLNLASKNSESISPSVALCIEB
NIPFINGSPONTFVPLIDPLAIKNSLIGDGFKSGQOTKMLVDVFLVGAIKPISI
VSNHIGNNDGMNLSAPQTFRSKEISKSNVVDVMSQNTLVEPGRPHVAVIKYVP
YVGDSCRANDEYTSIEFMGNTNTYVMNHCEDSLAAPITLIDVILAEISTLOKAB
BEDKRSRPRVATILSYLTKAPLVPPGTVPVNAISKORAMLENILRACVGLADENNMI
LEYK"

ORIGIN
      1.8%; Score 32; DB 8; Length 2053;
Query Match
Best Local Similarity 100.0%; Pred. No. 9.2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1233 GACAGCAAGAGCCATGATGATCACTTC 1264
|||||
1367 GACAGCAAGAGCCATGATGATCACTTC 1398

RESULT 11
AC125389/c
LOCUS AC125389 143012 bp DNA linear PLN 15-JUN-2004

```

```

DEFINITION Medicago truncatula clone mth2-12a18, complete sequence.
ACCESSION AC125389
VERSION AC125389.32 GI:48717535
KEYWORDS HTG.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 143012)
REFERENCE
AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-12a18
Unpublished
2 (bases 1 to 143012)
REFERENCE
AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (25-JUN-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 143012)
REFERENCE
AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (20-MAY-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 143012)
REFERENCE
AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
Submitted (15-JUN-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
ON Jun 15, 2004 this sequence version replaced gi:47524030.
COMMENT
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UDXNOR
-----
FEATURES
source
1.143012
Location/Qualifiers
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-12a18"
/clone_lib="Medicago truncatula BAC library H2"

ORIGIN
      1.8%; Score 32; DB 8; Length 143012;
Query Match
Best Local Similarity 100.0%; Pred. No. 7.2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1479 ACACCAAGTGTGATGATGATGTCGAAGACGCG 1510
|||||
13378 ACACCAAGTGTGATGATGATGTCGAAGACGCG 13347

RESULT 12
AY323824      1533 bp mRNA linear PLN 15-JUL-2003
LOCUS
DEFINITION Xerophyta viscosa myo-inositol-1-phosphate synthase INO1 mRNA,
complete cds.
ACCESSION AY323824
VERSION AY323824.1 GI:32492881
KEYWORDS
SOURCE Xerophyta viscosa
ORGANISM Xerophyta viscosa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Velloziaceae; Xerophyta.
1 (bases 1 to 1533)
REFERENCE

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:47:21 ; Search time 966.851 Seconds
(without alignments)
10775.958 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttcttctcttctgt.....aatgtatcataattctgtc 1760

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980a:*
2: geneseqn1990a:*
3: geneseqn2000a:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1760	100.0	1760	12	ADQ14490	Adq14490 Wild type
2	1760	100.0	1782	2	AAV62440	AAV62440 Soybean w
3	1533	87.1	1533	12	ADQ14498	Adq14498 Mutant so
4	1533	87.1	1533	13	AD881999	Ad881999 Soybean m
5	1533	87.1	1533	13	AD881993	Ad881993 Soybean m
6	1482	84.2	1533	2	AAV62443	AAV62443 Soybean m
7	1482	84.2	1533	13	ADQ14494	Adq14494 Mutant so
8	1482	84.2	1533	13	AD881997	Ad881997 Soybean m
9	146	8.3	1533	12	ADQ14504	Adq14504 Wild type
10	146	8.3	1533	12	ADQ14500	Adq14500 Mutant so
11	146	8.3	1533	12	ADQ14502	Adq14502 Mutant so
12	146	8.3	1533	13	AD882005	Ad882005 Soybean m
13	146	8.3	1533	13	AD882003	Ad882003 Soybean m
14	146	8.3	1533	13	AD882001	Ad882001 Soybean m
15	29	1.6	377	13	ADP93370	Adp93370 Cotton ex
16	29	1.6	566	13	ACN57874	Acn57874 Cotton gy
17	29	1.6	618	13	ACN62990	Acn62990 Cotton ca
18	26	1.5	1781	4	AAV62441	AAV62441 Soybean m
19	24	1.4	35	2	AAV62441	AAV62441 Soybean m
20	24	1.4	35	12	ADQ14492	Adq14492 Soybean m

21	23	1.3	345	12	ADP94497	Adp94497 Cotton ex
22	23	1.3	457	12	ADJ10863	Adj10863 Recombina
23	23	1.3	557	13	ACN58059	Acn58059 Cotton gy
24	23	1.3	601	13	ACN60643	Acn60643 Cotton gy
25	23	1.3	605	13	ACN60421	Acn60421 Cotton gy
26	23	1.3	1536	12	ADP43918	Adp43918 P. coarct
27	23	1.3	1578	2	AAV91656	AAV91656 Inositol-
28	23	1.3	1578	2	AAV65400	AAV65400 Inositol-
29	23	1.3	1665	2	AAV09006	AAV09006 Inducible
30	23	1.3	1931	2	AAV24407	AAV24407 Maize myo
31	23	1.3	1950	2	AAV90402	AAV90402 Nicotiana
32	23	1.3	3546	2	AAV24410	AAV24410 Maize myo
33	23	1.3	3546	2	AAV24411	AAV24411 Maize myo
34	21	1.2	35	13	AD881995	Ad881995 PCR prime
35	21	1.2	39	12	AAV62442	AAV62442 Soybean m
36	21	1.2	39	12	ADQ14493	Adq14493 Soybean m
37	20	1.1	257	12	ADP95252	Adp95252 Cotton ex
38	20	1.1	291	12	ADP93783	Adp93783 Cotton ex
39	20	1.1	388	13	ACN52468	Acn52468 Cotton an
40	20	1.1	389	13	ACN51365	Acn51365 Cotton an
41	20	1.1	449	12	ADP91072	Adp91072 Cotton ex
42	20	1.1	506	13	ACN57963	Acn57963 Cotton gy
43	20	1.1	512	10	AD881407	Ad881407 Arabidops
44	20	1.1	517	13	ACN63079	Acn63079 Cotton ca
45	20	1.1	525	13	ACN58143	Acn58143 Cotton gy

ALIGNMENTS

RESULT 1	ADQ14490	standard; cDNA, 1760 BP.
ID	ADQ14490	
XX	ADQ14490;	
AC		
XX		
DT	23-SEP-2004	(first entry)
XX		
DE	Wild type soybean myo-inositol 1-phosphate synthase cDNA #1.	
XX		
XX	Soybean; myo-inositol 1-phosphate synthase; gene; ss;	
KW	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;	
KM	inorganic phosphate.	
XX		
OS	Glycine max.	
XX		
FX		
HH		
FT	Key	Location/Qualifiers
FT	CDS	54..1586
FT		/tag= a
FT		/product= "Wild type soybean myo-inositol 1-phosphate synthase #1"
XX		
PN	US2004128713-A1.	
XX		
PD	01-JUL-2004.	
XX		
XX	21-NOV-2003; 2003US-00718952.	
XX		
PR	08-APR-1997; 97US-00835751.	
PR	07-APR-1998; 98WO-US006822.	
PR	26-APR-1999; 99US-00299315.	
PR	11-MAR-2002; 2002US-00025003.	
XX		
PA	(HITZ/) HITZ W D.	
PA	(SEBA/) SEBASTIAN S A.	
PA	(GRAC/) GRACE D J.	
PA	(STRE/) STREIT L G.	
XX		
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
XX		
DR	WPI; 2004-533135/51.	
DR	P-PsDB; ADQ14491.	
XX		

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX Claim 4; SEQ ID NO 1; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Query Match 100.0%; Score 1760; DB 12; Length 1760;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTTTTATTCCTTTGTAATTTTCAATTCATTCCTTAATCTTTGTAATTAATGTTCA 60
DB 1 CTTCTTTTATTCCTTTGTAATTTTCAATTCATTCCTTAATCTTTGTAATTAATGTTCA 60
QY 61 TCGAGAAATTTTAAGTTGAGTGTCTTAATGTGAAGTACA CCGAGCTGAATTCAGTCCG 120
DB 61 TCGAGAAATTTTAAGTTGAGTGTCTTAATGTGAAGTACA CCGAGCTGAATTCAGTCCG 120
QY 121 TGTACAACATCGAAGACCGAACTTGTTCACGAGAACGAAATGGCACTCATCACTGGA 180
DB 121 TGTACAACATCGAAGACCGAACTTGTTCACGAGAACGAAATGGCACTCATCACTGGA 180
QY 121 TGTACAACATCGAAGACCGAACTTGTTCACGAGAACGAAATGGCACTCATCACTGGA 180
DB 121 TGTACAACATCGAAGACCGAACTTGTTCACGAGAACGAAATGGCACTCATCACTGGA 180
QY 181 TTGTCAAAACCAATGTCTCAAAATAGAAATTTAAACCAACATCATGTTCTTAAATTAG 240
DB 181 TTGTCAAAACCAATGTCTCAAAATAGAAATTTAAACCAACATCATGTTCTTAAATTAG 240
QY 241 GGGTAAATGCTTGGGTGGGTGGGTAAGAACGAGCTCAACCTCAACGGTGTGTTATTG 300
DB 241 GGGTAAATGCTTGGGTGGGTGGGTAAGAACGAGCTCAACCTCAACGGTGTGTTATTG 300
QY 301 CTAAACGAGGAGGCAATTTCAATGGCTTACAAAGGACAAATTAATTAATTTG 360
DB 301 CTAAACGAGGAGGCAATTTCAATGGCTTACAAAGGACAAATTAATTAATTTG 360
QY 361 GGTCCCTCAACCAAGGCTCAGCTATCGAGTTGGGCTTTCCAGGGAGAGAAATCTATG 420
DB 361 GGTCCCTCAACCAAGGCTCAGCTATCGAGTTGGGCTTTCCAGGGAGAGAAATCTATG 420
QY 421 CCCCATTTCAAGAGCTGTTCCTCAATGTTAACTCTGACGACATTTGTTGGGGAGTGG 480
DB 421 CCCCATTTCAAGAGCTGTTCCTCAATGTTAACTCTGACGACATTTGTTGGGGAGTGG 480
QY 481 ATATTCAGCAATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTGTTGACATGATT 540
DB 481 ATATTCAGCAATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTGTTGACATGATT 540
QY 541 TGCAGAGAGAGTGAAGGCTTATAGGAATCCATGCTTCACTCCCGGAAATTAAGACC 600
DB 541 TGCAGAGAGAGTGAAGGCTTATAGGAATCCATGCTTCACTCCCGGAAATTAAGACC 600
QY 601 CGGATTTTCATGCTGCCAACCAAGAGAGCGTGCACAAACGTTCATCAAGGGCAAAAGC 660

DB 601 CGGATTTTCATGCTGCCAACCAAGAGAGCGTGCACAAACGTTCATCAAGGGCAAAAGC 660
QY 661 AAGACCAAGTTCAACAAATCATCAAGGCGGTTTAAAGAGCCACCAAGTGG 720
DB 661 AAGACCAAGTTCAACAAATCATCAAGGCGGTTTAAAGAGCCACCAAGTGG 720
QY 721 ACAAGTGGTGTACTGTGACCTGCCAACAGAGAGGTACGTAATTTGTTGGGCC 780
DB 721 ACAAGTGGTGTACTGTGACCTGCCAACAGAGAGGTACGTAATTTGTTGGGCC 780
QY 781 TTAATGACACATGAGAAATCTTTGGCTGTGTGACAGAAATGAGCTGAGATTTCTC 840
DB 781 TTAATGACACATGAGAAATCTTTGGCTGTGTGACAGAAATGAGCTGAGATTTCTC 840
QY 841 CTTCAACCTGTATATCCATTTGCTGTGTTATGSAATGTTCTTCAATATGAGAGC 900
DB 841 CTTCAACCTGTGTATCCATTTGCTGTGTTATGSAATGTTCTTCAATATGAGAGC 900
QY 901 CTCAAGACATTTTGTACAGGCTGATTTGATCTTGCATCGCAGAGAACATTTGATTG 960
DB 901 CTCAAGACATTTTGTATCAAGGCTGATTTGATCTTGCATCGCAGAGAACATTTGATTG 960
QY 961 GTGAGATGACTTCAAGAGTGTCAAGCCAAATGAATCTGTGTTGGTTCCTTG 1020
DB 961 GTGAGATGACTTCAAGAGTGTCAAGCCAAATGAATCTGTGTTGGTTCCTTG 1020
QY 1021 TGGGGGCTGTATCAAGGCAACATATATGATGATACCATCTGGGAAACAATGATG 1080
DB 1021 TGGGGGCTGTATCAAGGCAACATATATGATGATACCATCTGGGAAACAATGATG 1080
QY 1081 GTATGAATCTTTGGCTCCACAACTTTCCTTCCAGGAAATCTCAAGAGCAACGTTG 1140
DB 1081 GTATGAATCTTTGGCTCCACAACTTTCCTTCCAGGAAATCTCAAGAGCAACGTTG 1140
QY 1141 TTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAAATCCAGCATG 1200
DB 1141 TTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAAATCCAGCATG 1200
QY 1201 TTGTTGTTATTAAGTATGCTTACGTAGGGGACAGAGAGAGCAATGAGTGTGCA 1260
DB 1201 TTGTTGTTATTAAGTATGCTTACGTAGGGGACAGAGAGAGCAATGAGTGTGCA 1260
QY 1261 CTTCAAGATATTCATGGGTGAAAGACCACTTTGTTGCAACATGCGAGATT 1320
DB 1261 CTTCAAGATATTCATGGGTGAAAGACCACTTTGTTGCAACATGCGAGATT 1320
QY 1321 CCCTTTAGCTGCTCTATTTATCTTGACCTTGCTCTTGTGAGCTCAGACTAGAA 1380
DB 1321 CCCTTTAGCTGCTCTATTTATCTTGACCTTGCTCTTGTGAGCTCAGACTAGAA 1380
QY 1381 TCGAGTTTAAACCTGAAATAGGGGAAATTCACATTCACCAAGTGTCTACATCC 1440
DB 1381 TCGAGTTTAAACCTGAAATAGGGGAAATTCACATTCACCAAGTGTCTACATCC 1440
QY 1441 TCAGTACTCTCAACAGGCTCTCTGTTCACACCGGGTACACAGTGTGAAATGATGT 1500
DB 1441 TCAGTACTCTCAACAGGCTCTCTGTTCACACCGGGTACACAGTGTGAAATGATGT 1500
QY 1501 CAAAGACGCTGCAATGCTGGAATAATATGAGGCTTTGTGTGATTGGCCCAAGAGA 1560
DB 1501 CAAAGACGCTGCAATGCTGGAATAATATGAGGCTTTGTGTGATTGGCCCAAGAGA 1560
QY 1561 ATAAATGATTTCTGAGTACAAAGTAAAGATGGGACCGAAGATTAATGTTGGGGTGA 1620
DB 1561 ATAAATGATTTCTGAGTACAAAGTAAAGATGGGACCGAAGATTAATGTTGGGGTGA 1620
QY 1621 CCTAGCTGAATTTTATGTTAATATATGTTGCTTAATTTTGAAGTAAATTTGA 1680
DB 1621 CCTAGCTGAATTTTATGTTAATATATGTTGCTTAATTTTGAAGTAAATTTGA 1680
QY 1681 TGCATCAGCTTCAATTAATGCTTTAAGCGGGCATATTTCTGTTTCTAGAGAAATGATG 1740


```

QY 1141 TTGATGATATGTCAGACCAATGCAATCTCTATGAGCCTGCTGAACATCCAGACCATG 1200
DB 1141 TTGATGATATGTCAGACCAATGCAATCTCTATGAGCCTGCTGAACATCCAGACCATG 1200
QY 1201 TTGTTGTTATTAAGTATGTCCTTACGTAGGGGAGACGAAGAAGCCATGATGATGATA 1260
DB 1201 TTGTTGTTATTAAGTATGTCCTTACGTAGGGGAGACGAAGAAGCCATGATGATGATA 1260
QY 1261 CTTCAGAGATATTCATGCGGTGAGAAAGACCAATGTTTGTGCAACAACATCCAGAGATT 1320
DB 1261 CTTCAGAGATATTCATGCGGTGAGAAAGACCAATGTTTGTGCAACAACATCCAGAGATT 1320
QY 1321 CCTCTTGTAGCTGCTCTATATATCTTGTGACTTGTCTTCTTGTGAGCTCAGCACTAGAA 1380
DB 1321 CCTCTTGTAGCTGCTCTATATATCTTGTGACTTGTCTTCTTGTGAGCTCAGCACTAGAA 1380
QY 1381 TCGAGTTTAAAGCTGAAATGAGGAGAAATTCATCTCATTCACACCACTGCTACCATCC 1440
DB 1381 TCGAGTTTAAAGCTGAAATGAGGAGAAATTCATCTCATTCACACCACTGCTACCATCC 1440
QY 1441 TCGAGTACCTCAGCAAGGCTCTCTGTGTTCCACCGGGTACACCACTGATGATGATTGT 1500
DB 1441 TCGAGTACCTCAGCAAGGCTCTCTGTGTTCCACCGGGTACACCACTGATGATGATTGT 1500
QY 1501 CAAAGCAGCGCTGCAATGCTGAGAAACATATGAGGGCTTGTGTTGATTGSCCCAGAGA 1560
DB 1501 CAAAGCAGCGCTGCAATGCTGAGAAACATATGAGGGCTTGTGTTGATTGSCCCAGAGA 1560
QY 1561 ATAACTAGATCTCTGAGTACAGAGATGAGGAGATGAGGAGAAATATATATGTTGGGTAG 1620
DB 1561 ATAACTAGATCTCTGAGTACAGAGATGAGGAGATGAGGAGAAATATATATGTTGGGTAG 1620
QY 1621 CTGAGTGAATGTTTATGTTATATATATATGTTTGTGCTTATATTTGCAAGTATATGAA 1680
DB 1621 CTGAGTGAATGTTTATGTTATATATATATGTTTGTGCTTATATTTGCAAGTATATGAA 1680
QY 1681 TGATGAGCTTCATATATGCTTTAGACGGGGGCAATCTGTTTACTAGGAACATGATG 1740
DB 1681 TGATGAGCTTCATATATGCTTTAGACGGGGGCAATCTGTTTACTAGGAACATGATG 1740
QY 1741 AATGATGATATATTTTGTGT 1760
DB 1741 AATGATGATATATTTTGTGT 1760

```

RESULT 3
ADQ14498
ID ADQ14498 strand; cDNA; 1533 BP.
XX
AC ADQ14498;
DT 23-SEP-2004 (first entry)
XX
DB Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.
XX
KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
OS Glycine max.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #2"
XX
PD US2004128713-A1.
XX 01-JUL-2004.

```

PF 21-NOV-2003; 2003US-00718952.
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98MO-US006822.
PR 26-APR-1999; 99US-00293315.
PR 11-MAR-2002; 2002US-00025003.
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
DR WPI; 2004-533135/51.
DR P-PSDB; ADQ14499.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Example 8; SEQ ID NO 9; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
SQ

```

Query Match 87.1%; Score 1533; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 54 ATGTTCAATCGAGAAATTTTAAGGTTGAGTCTCTAATGTAAGTACACCGAGACTGAGATT 113
DB 1 ATGTTCAATCGAGAAATTTTAAGGTTGAGTCTCTAATGTAAGTACACCGAGACTGAGATT 60
QY 114 CAGTCCGCTGATCAATTAAGAAACCAACGAACTTGTTCACGAGACAGAAATGCACTAT 173
DB 61 CAGTCCGCTGATCAATTAAGAAACCAACGAACTTGTTCACGAGACAGAAATGCACTAT 120
QY 174 CAGTGGATTGCAAAACCAAAATCTGTCAAAATGCAAAATTTAAACCAACATCATGTTCTT 233
DB 121 CAGTGGATTGCAAAACCAAAATCTGTCAAAATGCAAAATTTAAACCAACATCATGTTCTT 180
QY 234 AAATTAGGGGTAAATGCTTGTGGGTGGGAGAAACAGGCTCAACCTCAACGGGTGT 293
DB 181 AAATTAGGGGTAAATGCTTGTGGGTGGGAGAAACAGGCTCAACCTCAACGGGTGT 240
QY 294 GTTATGCTAACGAGAGGGGATTTCAATGGGCTTCAAAAGACCAAGATTTCAACAGCCAT 353
DB 241 GTTATGCTAACGAGAGGGGATTTCAATGGGCTTCAAAAGACCAAGATTTCAACAGCCAT 300
QY 354 TACTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 413
DB 301 TACTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 360

```

```

OY 414 ATCTATGCCCCCATTCAGAGAGCTGCTTCCAGTGAATTAACCTGACGACATGTGTGGG 473
DB 361 ATCTATGCCCCCATTCAGAGAGCTGCTTCCAGTGAATTAACCTGACGACATGTGTGGG 420
OY 474 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAAAGTGTGGAC 533
DB 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAAAGTGTGGAC 480
OY 534 ATCGATTTTGCAGAGCAGCTTGAAGGCTTACATGGAATCCATGCTTCCATCCCGGAATC 593
DB 481 ATCGATTTTGCAGAGCAGCTTGAAGGCTTACATGGAATCCATGCTTCCATCCCGGAATC 540
OY 594 TATGACCCGGATTTCAATGCTGCTCCCAACAGAGAGGCTGCCAACAACGTCATCAAGGGC 653
DB 541 TATGACCCGGATTTCAATGCTGCTCCCAACAGAGAGGCTGCCAACAACGTCATCAAGGGC 600
OY 654 ACAAGCAAGAGCAGCTTCAACAAATCATCAAGAGATCAAGGCGTTTAAAGAGCCACC 713
DB 601 ACAAGCAAGAGCAGCTTCAACAAATCATCAAGAGATCAAGGCGTTTAAAGAGCCACC 660
OY 714 AAAGTGCAACAGGTGTGTGACTGTGACCTGCCAACAAGAGAGTACAGTAATTTGGTT 773
DB 661 AAAGTGCAACAGGTGTGTGACTGTGACCTGCCAACAAGAGAGTACAGTAATTTGGTT 720
OY 774 GTGGGCGCTTAATGACACATGAGAAATCTTGGCTGTGTGAGCAAGAAATGAGGCTGAG 833
DB 721 GTGGGCGCTTAATGACACATGAGAAATCTTGGCTGTGTGAGCAAGAAATGAGGCTGAG 780
OY 834 ATTTCTCTCTCCACCTTGTATGTCATGCTTGTGTATGGAATAATGCTCTTTCATTAAT 893
DB 781 ATTTCTCTCTCCACCTTGTATGTCATGCTTGTGTATGGAATAATGCTCTTTCATTAAT 840
OY 894 GGAAGCCCTCAGAAACCTTTTGTACCAAGGCTGATTTGATTTGCCATCCGAGAAACACT 953
DB 841 GGAAGCCCTCAGAAACCTTTTGTACCAAGGCTGATTTGATTTGCCATCCGAGAAACACT 900
OY 954 TTGATTTGGTGAAGATGACTTCAAGAGTGGTCAAGCAACAAATGAATACTGTGTGTGAT 1013
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGGTCAAGCAACAAATGAATACTGTGTGTGAT 960
OY 1014 TTCTCTTGGGGGCTGTGTATCAAGCCAACTTATAGTCAGTTTCAACCACTTGGGAAAC 1073
DB 961 TTCTCTTGGGGGCTGTGTATCAAGCCAACTTATAGTCAGTTTCAACCACTTGGGAAAC 1020
OY 1074 AATGATGTATGATATTTTGGCTCCAAACCTTTCCGTTCCAAAGAAATCTCCAAAGAC 1133
DB 1021 AATGATGTATGATATTTTGGCTCCAAACCTTTCCGTTCCAAAGAAATCTCCAAAGAC 1080
OY 1134 AACGTTGTGATGATATGTCACACCAATGCAATCTCTATGAGGCTGTGAACATCCA 1193
DB 1081 AACGTTGTGATGATATGTCACACCAATGCAATCTCTATGAGGCTGTGAACATCCA 1140
OY 1194 GACCATGTTGTTTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGCATGAT 1253
DB 1141 GACCATGTTGTTTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGCATGAT 1200
OY 1254 GAGTACACTTCAGAGATTTTCATGSGTGAAGAGACCAATGTTTTGCAACAACATGC 1313
DB 1201 GAGTACACTTCAGAGATTTTCATGSGTGAAGAGACCAATGTTTTGCAACAACATGC 1260
OY 1314 GAGGATTCCTCTTACTGCTCTCTATTAATCTTGAATGCTGCTTCTTGTGCTGAC 1373
DB 1261 GAGGATTCCTCTTACTGCTCTCTATTAATCTTGAATGCTGCTTCTTGTGCTGAC 1320
OY 1374 ACTAGATGAGATTTAAAGCTGAATAATGAGGAAATTCACACTATTCACACCAAGTTGCT 1433
DB 1321 ACTAGATGAGATTTAAAGCTGAATAATGAGGAAATTCACACTATTCACACCAAGTTGCT 1380
OY 1434 ACCATCTCACTCACTCAACAAAGGCTCTCTGCTTCCACCGGGTACACCAAGTGTGAAT 1493
DB 1381 ACCATCTCACTCACTCAACAAAGGCTCTCTGCTTCCACCGGGTACACCAAGTGTGAAT 1440
OY 1494 GCATTGTCAAAAGAGGCTGCAATGCTGAAACATTAATAGGCTGTGTGTGATTTGGCC 1553

```

```

DB 1441 GCATTGTCAAAAGAGGCTGCAATGCTGGAATAATATAGGCTGTGTGTGATTTGGCC 1500
OY 1554 CCAGAGATTAACATGATTTCTCGATGACAACTGA 1586
DB 1501 CCAGAGATTAACATGATTTCTCGATGACAACTGA 1533

RESULT 4
ADS81999
ID ADS81999 standard; cDNA; 1533 BP.
XX
AC ADS81999;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29004JP01.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
PI Hitz WD, Sebastian SA;
XX
XX MPI: 2004-639957/62.
XX P-PSDB; ADS82000.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Example 8; SEQ ID NO 9; 34bp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis of myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g), provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived
XX from seeds of a soybean plant (homozygous for one or more gene encoding a
XX mutant myo-inositol 1-phosphate synthase having decreased capacity for
XX the synthesis of myo-inositol 1-phosphate, where the gene confers a
XX heritable phenotype as mentioned above), and making or producing a

```

CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.
XX

Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Query Match 87.1%; Score 1533; DB 13; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 54 ATGTCATGAGAAATTTTAAAGTTGAGTCTCTAATGTGAATACCGAGCTGAGAT 113
DB 1 ATGTCATGAGAAATTTTAAAGTTGAGTCTCTAATGTGAATACCGAGCTGAGAT 60
OY 114 CAGTCGGTGTACAACTACGAAACCAACCGAATCTTGTACGAGAACGAAATGGCACTAT 173
DB 61 CAGTCGGTGTACAACTACGAAACCAACCGAATCTTGTACGAGAACGAAATGGCACTAT 120
OY 174 CAGTGGATTGTCAAAACCCAAATCTGTCAATACGAAATTTAAACCAACATCATGTTCT 233
DB 121 CAGTGGATTGTCAAAACCCAAATCTGTCAATACGAAATTTAAACCAACATCATGTTCT 180
OY 234 AAATTAAGGGGTAAATGTTGGGTTGGGGTGGAAACAGGGCTCAACCTCACCGGTGT 293
DB 181 AAATTAAGGGGTAAATGTTGGGTTGGGGTGGAAACAGGGCTCAACCTCACCGGTGT 240
OY 294 GTTATTTGCTAACCGAGAGGGCAATTCATGGGCTACAAAGACAAAGATTCAAGCAAT 353
DB 241 GTTATTTGCTAACCGAGAGGGCAATTCATGGGCTACAAAGACAAAGATTCAAGCAAT 300
OY 354 TACTTTGGCTCCCTCAACCCAAAGCTCAGTATCCAGTTGGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAAAGCTCAGTATCCAGTTGGGTCTTCCAGGAGAGAA 360
OY 414 ATCTAATGCCCAATTCAGAGAGCTGCTTCAATGTGTTAACCTGACGACAAATGTTGGG 473
DB 361 ATCTAATGCCCAATTCAGAGAGCTGCTTCAATGTGTTAACCTGACGACAAATGTTGGG 420
OY 474 GGATGGATATACGAACATGAACCTGGCTGATGCGATGGCCAGGGCAAAAGGTGTTGAC 533
DB 421 GGATGGATATACGAACATGAACCTGGCTGATGCGATGGCCAGGGCAAAAGGTGTTGAC 480
OY 534 ATCGATTGCAAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 593
DB 481 ATCGATTGCAAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 540
OY 594 TATGACCCCGGATTTCTTGTCTGCCAACAAAGAGACGGTCCAAACGTCATCAAGGGC 653
DB 541 TATGACCCCGGATTTCTTGTCTGCCAACAAAGAGACGGTCCAAACGTCATCAAGGGC 600
OY 654 ACAAGCAAGAGCAAGTTCAACAAATCATCAAGATCAAGGCGTTTAAAGAACCAAC 713
DB 601 ACAAGCAAGAGCAAGTTCAACAAATCATCAAGATCAAGGCGTTTAAAGAACCAAC 660
OY 714 AAAGTGACAAAGGTGTTGTTACTGTGACTGCAACACAGAGGTTACAGTAATTTGGTT 773
DB 661 AAAGTGACAAAGGTGTTGTTACTGTGACTGCAACACAGAGGTTACAGTAATTTGGTT 720
OY 774 GTGGGCTTAATACACCATGAGGAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
DB 721 GTGGGCTTAATACACCATGAGGAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
OY 834 ATTTCTCTTCCACTGTATGCAATGCTGTGTATGGAATAATGTCCTTCAATTAAT 893
DB 781 ATTTCTCTTCCACTGTATGCAATGCTGTGTATGGAATAATGTCCTTCAATTAAT 840
OY 894 GGAAGCCCTCAGAAACCTTTGTACAGAGGCTGATTTGTCATCGGAGAAACACT 953
DB 841 GGAAGCCCTCAGAAACCTTTGTACAGAGGCTGATTTGTCATCGGAGAAACACT 900
```

```

DB 841 GGAAGCCCTCAGAAACCTTTGTACAGAGGCTGATTTGTCATCGGAGAAACACT 900
OY 954 TTGATTTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGGTAT 1013
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGGTAT 960
OY 1014 TTCTTTGGGGGCTGTATCAAGCCAACTATATAGTCAATTACACCATCTGGGAAC 1073
DB 961 TTCTTTGGGGGCTGTATCAAGCCAACTATATAGTCAATTACACCATCTGGGAAC 1020
OY 1074 AATGATGTGAAGATCTTGGGCTCCAAACCTTCCGTTCCAGGAATCTCCAAAGAC 1133
DB 1021 AATGATGTGAAGATCTTGGGCTCCAAACCTTCCGTTCCAGGAATCTCCAAAGAC 1080
OY 1134 AACGTTGTGATGATATGTCACAGCAATGCATCTCTATAGACCTGTGAACATCCA 1193
DB 1081 AACGTTGTGATGATATGTCACAGCAATGCATCTCTATAGACCTGTGAACATCCA 1140
OY 1194 GACCATGTTGTTGTTATTAATGATGTGCTTACGTAAGGGGACAGCAAGAGCCATGAT 1253
DB 1141 GACCATGTTGTTGTTATTAATGATGTGCTTACGTAAGGGGACAGCAAGAGCCATGAT 1200
OY 1254 GAGTACCTTCAAGATATTTCAAGGGTGAAGAGACACATTTGTTGACAAACATGAC 1313
DB 1201 GAGTACCTTCAAGATATTTCAAGGGTGAAGAGACACATTTGTTGACAAACATGAC 1260
OY 1314 GAGATTCCTCTTATGCTCTCTATTAATCTTGAATGTTGCTTCTTGTGAGCTCAGC 1373
DB 1261 GAGATTCCTCTTATGCTCTCTATTAATCTTGAATGTTGCTTCTTGTGAGCTCAGC 1320
OY 1374 ACTGAATGAGATTTAAAGCTGAATAATGAGGAAATTCACATCTTCCACCAATGCT 1433
DB 1321 ACTGAATGAGATTTAAAGCTGAATAATGAGGAAATTCACATCTTCCACCAATGCT 1380
OY 1434 ACCATCTCAGCTACCTCAACCAAGGCTCCTGTTCCACGGGATACCAAGTGTGAT 1493
DB 1381 ACCATCTCAGCTACCTCAACCAAGGCTCCTGTTCCACGGGATACCAAGTGTGAT 1440
OY 1494 GCATTGTCAAAAGCAGCTGCAATGCTGGAATAACATATGAGGGCTTGTGTGATGGC 1553
DB 1441 GCATTGTCAAAAGCAGCTGCAATGCTGGAATAACATATGAGGGCTTGTGTGATGGC 1500
OY 1554 CCAGAGATTAACATGATTTCTGAGTACAAATGA 1586
DB 1501 CCAGAGATTAACATGATTTCTGAGTACAAATGA 1533

RESULT 5
ADS81993
ID ADS81993 standard; cDNA; 1533 BP.
XX
AC ADS81993;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
```

PR	08-APR-1997;	97US-00835751.
PR	07-APR-1998;	98MO-US006822.
XX		
PA	(HITZ/) HITZ W D.	
PA	(SEBA/) SEBASTIAN S A.	
XX		
PL	Hitz WD, Sebastian SA;	
XX		
DR	WP1; 2004-639957/62.	
DR	P-PSDB; ADS81994.	
XX		
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.	
XX		
PS	Claim 2; SEQ ID NO 1; 34p; English.	
CC		
CC	The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homologous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.	
CC		
XX		
XX	Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;	
XX		
QY	Query Match	87.1%; Score 1533; DB 13; Length 1533;
QY	Best Local Similarity	100.0%; Pred. No. 0;
Db	Matches 1533; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	54 ATGTTTCATCGAGAATTTTAAAGTTGAGTGTCTTAAATGTAAGTACCGAGACTGAGATT	113
Db	1 ATGTTTCATCGAGAATTTTAAAGTTGAGTGTCTTAAATGTAAGTACCGAGACTGAGATT	60
QY	114 CAGTCCGGTGTACACTACGAAACCCGGAATTGTTACAGAGAACGAGATGGCACTTAT	173
Db	61 CAGTCCGGTGTACACTACGAAACCCGGAATTGTTACAGAGAACGAGATGGCACTTAT	120
QY	174 CAGTGGATGTGCAAAACCCAAATCTGTCAAAATTCGAAATTTAAACCAATCGATGTTCTT	233
Db	121 CAGTGGATGTGCAAAACCCAAATCTGTCAAAATTCGAAATTTAAACCAATCGATGTTCTT	180
QY	234 AAATTAGGGGTAAATGCTTGTGGGTGGGGTGGAAACAACGGCTCAACCTCAACGGTGGT	293
Db	181 AAATTAGGGGTAAATGCTTGTGGGTGGGGTGGAAACAACGGCTCAACCTCAACGGTGGT	240
QY	294 GTTATTTGTTAAACCGAGAGGGCATTTTCATATGGGCTTCAAAAGACAGATTTCACAGCCAAAT	353
Db	241 GTTATTTGTTAAACCGAGAGGGCATTTTCATATGGGCTTCAAAAGACAGATTTCACAGCCAAAT	300

OY	354	TACTTTGGCTCCCTCACCCAAAGCTCAGCTATCCGAGTGGGTCTTCCAGGGAGAGAA	413
Db	301	TACTTTGGCTCCCTCACCCAAAGCTCAGCTATCCGAGTGGGTCTTCCAGGGAGAGAA	360
OY	414	ATCTATGGCCCACTTCAAGAGCCTCTTCCAAATGSGTTAAACCTGACGACATGTGTGGG	473
Db	361	ATCTATGGCCCACTTCAAGAGCCTCTTCCAAATGSGTTAAACCTGACGACATGTGTGGG	420
OY	474	GGATGGATATCAGCAACATGACCTGAGTATGCCATGCGCAGGGCAAGAGTGTGGAC	533
Db	421	GGATGGATATCAGCAACATGACCTGAGTATGCCATGCGCAGGGCAAGAGTGTGGAC	480
OY	534	ATCGATTTGCAAGAGCAGTTGAGGCTTTACATGGAATTCATGCTTCCACTCTCCCGAATC	593
Db	481	ATCGATTTGCAAGAGCAGTTGAGGCTTTACATGGAATTCATGCTTCCACTCTCCCGAATC	540
OY	594	TATGACCCGGATTTCATTTGCTGCCAACAGAGAGGGGTGCACAAACGTATCAAGGGC	653
Db	541	TATGACCCGGATTTCATTTGCTGCCAACAGAGAGGGGTGCACAAACGTATCAAGGGC	600
OY	654	ACAAGACAAGAGCAAGTTCAAACATATCAACAAACATCAAGGCGTTTAAAGAACCCAC	713
Db	601	ACAAGACAAGAGCAAGTTCAAACATATCAACAAACATCAAGGCGTTTAAAGAACCCAC	660
OY	714	AAAGTGCACAAAGTGGTGTATCTGTGACCTGCCAACAGAGAGGTACAGTAATTTGGTT	773
Db	661	AAAGTGCACAAAGTGGTGTATCTGTGACCTGCCAACAGAGAGGTACAGTAATTTGGTT	720
OY	774	GTGGGCTTAAATGACACCATGAGAAATCTCTTGGCTGCTGTGACAGAAATGAGGCTGAG	833
Db	721	GTGGGCTTAAATGACACCATGAGAAATCTCTTGGCTGCTGTGACAGAAATGAGGCTGAG	780
OY	834	ATTTCTCCTTCCACCTTGTAATGACATGCTGTGTATGAGAAATGTCTCTTCAATTAAT	893
Db	781	ATTTCTCCTTCCACCTTGTAATGACATGCTGTGTATGAGAAATGTCTCTTCAATTAAT	840
OY	894	GGAAGCCCTCAGAACACTTTTGTATCCAGGGCTATTTGATCTTGCATCGCAGGAAACAT	953
Db	841	GGAAGCCCTCAGAACACTTTTGTATCCAGGGCTATTTGATCTTGCATCGCAGGAAACAT	900
OY	954	TTGATTTGCTGAGATGACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTGTGGTAT	1013
Db	901	TTGATTTGCTGAGATGACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTGTGGTAT	960
OY	1014	TTCTTGTGGGGCTGGTATCAAGCAACATCTATAGCTGTTCAACACATCTGGGAAAC	1073
Db	961	TTCTTGTGGGGCTGGTATCAAGCAACATCTATAGCTGTTCAACACATCTGGGAAAC	1020
OY	1074	AATGATGATATGATCTTTGGGCTCCAAACATTTCCGTTCCAGAGAAATCTCCMAAGC	1133
Db	1021	AATGATGATATGATCTTTGGGCTCCAAACATTTCCGTTCCAGAGAAATCTCCMAAGC	1080
OY	1134	AACGTTGTATGATATGATCAACAGCAATGCCATCTATGAGCCTGGTGAACATCA	1193
Db	1081	AACGTTGTATGATATGATCAACAGCAATGCCATCTATGAGCCTGGTGAACATCA	1140
OY	1194	GACCATTTGTGTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGCCATGAT	1253
Db	1141	GACCATTTGTGTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGCCATGAT	1200
OY	1254	GAGTACACTTCAGAGATATTCATGGGTGAAAGAGCACATGTTGTTTGCAACACATGAC	1313
Db	1201	GAGTACACTTCAGAGATATTCATGGGTGAAAGAGCACATGTTGTTTGCAACACATGAC	1260
OY	1314	GAGGATTCCTCTTACGTGCTCCTATTAATCTTGGACTTGGTCTTCTTGCTGAGCTCAGC	1373
Db	1261	GAGGATTCCTCTTACGTGCTCCTATTAATCTTGGACTTGGTCTTCTTGCTGAGCTCAGC	1320
OY	1374	ACTAGATTCGAGTTTAAAGCTGAAATGAGGAAATTCACATCATCTCCACCCAGTTGCT	1433
Db	1321	ACTAGATTCGAGTTTAAAGCTGAAATGAGGAAATTCACATCATCTCCACCCAGTTGCT	1380

QY 1434 ACCATCTCAGCTACTCTACCAAGGCTCCTCTGTTCCACCGGTAACACAGTGTGAT 1493
DB 1381 ACCATCTCAGCTACTCTACCAAGGCTCCTCTGTTCCACCGGTAACACAGTGTGAT 1440
QY 1494 GCATTGTCAAAGCAGCTGCAATGCTGAAAACATPATAGGGCTTGTGTGATTGGCC 1553
DB 1441 GCATTGTCAAAGCAGCTGCAATGCTGAAAACATPATAGGGCTTGTGTGATTGGCC 1500
QY 1554 CCAGAGAAATACATGATTTCTCGAGTACAGTGA 1586
DB 1501 CCAGAGAAATACATGATTTCTCGAGTACAGTGA 1533

RESULT 6
AAV62443
ID AAV62443 standard; cDNA; 1533 BP.

AC AAV62443;
XX
XX 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
XX Soybean mutant myo-inositol 1-phosphate synthase cDNA.

XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid; de.
XX

OS Glycine max; line LR33.

PN W09845448-A1.

PD 15-OCT-1998.

PF 07-APR-1998; 98WO-US006822.

PR 08-APR-1997; 97US-00835751.

PA (DUPO) DU PONT DE NEMOURS & CO E. I.

PI Hitz WD, Sebastian SA;

XX WPI; 1998-568353/48.

DR P-PSDB; AAW79741.

PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.

PS Example 5; Page 48-49; 63p; English.

XX This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
XX inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
CC clone was isolated from a cDNA library of soybean line LR33 by PCR
CC amplification (see AAV62441-42). Line LR33 was obtained by chemical
CC mutagenesis of wild-type soybean genome and as a reduced raffinose
CC saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. Sequencing revealed a single base change
CC mutation (G to T at base 1241) in the LR33 sequence when compared to the
CC wild-type sequence (see AAV62440). The mutation results in a seed
CC phenotype of very low raffinose saccharide sugars, very high sucrose and
CC low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
CC raffinose saccharide, sucrose, phytic acid and inorganic phosphate
CC content of soybean seeds, leading to useful soybean products, e.g. a seed
CC phytic acid content of less than 17 ug/g, a seed content of raffinose and
CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 84.2%; Score 1482; DB 2; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 ATGTCATCGAATATTTTAAAGTTGAGTGTCTTAATGTAATGTAACCGAGACTGAGATT 113
DB 1 ATGTCATCGAATATTTTAAAGTTGAGTGTCTTAATGTAATGTAACCGAGACTGAGATT 60
QY 114 CAGTCCGTGTACACTAGAAAACCGAAGCTTGTTCACGAGAACAGAAATGGCACTAT 173
DB 61 CAGTCCGTGTACACTAGAAAACCGAAGCTTGTTCACGAGAACAGAAATGGCACTAT 120
QY 174 CAGTGGATTGTCAACCCCAATCTGTCAAAATACGAATTTAAACCAATCCATGTTCT 233
DB 121 CAGTGGATTGTCAACCCCAATCTGTCAAAATACGAATTTAAACCAATCCATGTTCT 180
QY 234 AAATTAGGGGTAAATGCTTGTGGGTGGGAAACAACGCTCAACCTCAACGGTGT 293
DB 181 AAATTAGGGGTAAATGCTTGTGGGTGGGAAACAACGCTCAACCTCAACGGTGT 240
QY 294 GTTATTGCTAACCGAGAGGCAATTCATGGGCTACAAAGGACAGATTCAACAGCCAT 353
DB 241 GTTATTGCTAACCGAGAGGCAATTCATGGGCTACAAAGGACAGATTCAACAGCCAT 300
QY 354 TACTTGGGCTCCCTCACCACCAAGCTCAGCTATCCGAGTTGGGCTCTTCAGGAGAGAA 413
DB 301 TACTTGGGCTCCCTCACCACCAAGCTCAGCTATCCGAGTTGGGCTCTTCAGGAGAGAA 360
QY 414 ATCTATGCCCAATTCAGAGGCTGCTTCAATGTTAAACCTGACGACATTGTGTGG 473
DB 361 ATCTATGCCCAATTCAGAGGCTGCTTCAATGTTAAACCTGACGACATTGTGTGG 420
QY 474 GGATGGGATATCAGCAACATGAACCTGGCTGATGGCAAGGCCAAAGGTGTTTAC 533
DB 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGGCAAGGCCAAAGGTGTTTAC 480
QY 534 ATCGATTTGACAGAACGATTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 593
DB 481 ATCGATTTGACAGAACGATTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 540
QY 594 TATGACCCGGATTTTATTGCTGCGCAACAAGAGAGCGTCCCAACAGCTCATCAAGG 653
DB 541 TATGACCCGGATTTTATTGCTGCGCAACAAGAGAGCGTCCCAACAGCTCATCAAGG 600
QY 654 ACNAAAGCAAGCAAGTTCAACAAATCATCAAGATCAAGAGCGTTAAAGAACCAAC 713
DB 601 ACNAAAGCAAGCAAGTTCAACAAATCATCAAGATCAAGAGCGTTAAAGAACCAAC 660
QY 714 AAAGTGACAAAGGCTGTTGATCTGTGACTGCACTGCAACAGAGAGGTACATATTGGTT 773
DB 661 AAAGTGACAAAGGCTGTTGATCTGTGACTGCACTGCAACAGAGAGGTACATATTGGTT 720
QY 774 GTGGGCTTTAATGACCAATGAGATCTTTGCTGCTGTGACAGAAATGAGGCTGAG 833
DB 721 GTGGGCTTTAATGACCAATGAGATCTTTGCTGCTGTGAGACAGAAATGAGGCTGAG 780
QY 834 ATTTTCTCTTCCACTTGTATGCCATTTGCTGTGTATGAAATGTTCTTTCAATTAAT 893
DB 781 ATTTTCTCTTCCACTTGTATGCCATTTGCTGTGTATGAAATGTTCTTTCAATTAAT 840
QY 894 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAACT 953
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAACT 900
QY 954 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGAT 1013
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGAT 960
QY 1014 TTCTTTTGGGGGCTGTGATCAAGCAACATTAATGTCAGTTAAACAATTTGGGAAC 1073
DB 961 TTCTTTTGGGGGCTGTGATCAAGCAACATTAATGTCAGTTAAACAATTTGGGAAC 1020
QY 1074 AATGATGATGATATCTTGGGCTCCAAACACTTTCGTTCCAGAGAAATCTCAAGAGC 1133
DB 1021 AATGATGATGATATCTTGGGCTCCAAACACTTTCGTTCCAGAGAAATCTCAAGAGC 1080
QY 1134 AACGTTTGTGATGATGATGATCAACAGCAATGCAATCTCTATGAGCTGTGGAACATCCA 1193

AUTHORS Majee, M., Majumder, A.N.L. and Mundree, S.G.
TITLE Molecular characterization of Xynol, a myo-inositol-1-phosphate synthase from the resurrection plant *Xerophyta vascosa*
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1533)
AUTHORS Majee, M., Majumder, A.N.L. and Mundree, S.G.
TITLE Direct Submissions
JOURNAL Submitted (14-JUN-2003) Molecular and Cell Biology, University of Cape Town, Private Bag, Rondebosch, Western Cape 7701, South Africa
FEATURES Location/Qualifiers
 1.1533
 /organism="Xerophyta vascosa"
 /mol_type="mRNA"
 /db_xref="taxon:90708"
 1.1533
 /note="Xynol"
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase INO1"
 /protein_id="AA085531.1"
 /db_xref="GI:32492882"
 /translation="MTESFRVSPHYRGAAIESDYDTTELTVSHSDGASRVV
 RPKSVNFRRTTTPKLGVMVGWGNNGSTLTAGVIANRGI SWATDKYQOANY
 GSLTQASTIRVGSYNGEIEIYAPFKSLIPVNPDLVFGMDISNNLADAMRAKYLD
 IDLOKQRLPYMESMVLPGLIYDPDLFIAANGSRANVIGTKKEQEOIIOIRBEKE
 KSKVDKVVYLVMTANTERSNVYGVNDTENTMILASVDKRAEISPTIYAIAVCVMEGI
 PTNGSPQNTFVPGILDLAIKKNCLIGGDFPSGQTKKSVLVDPLVGAIGKPTSVS
 VNLGNNDGMNLSAPQTFRSKEISNSVDDVANGSLIYEPGHPDHVVVIVKYVCV
 GDSKRAMDEYTSIEIFMGSKNTIVMNTCEDSLAIPILIDVLVLAELTRIDFMSENE
 EKFSHFHPVATILSYLTKAPLVPGTPVNVNALAKORAMLENIVRACVGLAPENNMILE
 YK"

ORIGIN
 Query Match 1.5%; Score 26; DB 8; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 AGCAATGCATCCTCATAGACCTGG 1183
 1105 AGCAATGCATCCTCATAGACCTGG 1130

RESULT 13
LOCUS AY096554 1564 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana putative myo-inositol-1-phosphate synthase (At5g10170) mRNA, complete cds.
ACCESSION AY096554
VERSION AY096554.1 GI:20465488
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1564)
 Yamada, K., Bahh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, J.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carinci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Open Reading Frame (ORF) Clones
 Unpublished
 2 (bases 1 to 1564)
 Yamada, K., Bahh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, J.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carinci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,

TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carinci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.
 The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Bahh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, J.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
 Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.
 Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.
FEATURES Location/Qualifiers
 1.1564
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="U018590"
 /ecotype="Columbia"
 /note="This clone is in pUNI 51."
 1.1564
 /gene="At5g10170"
 1.1533
 /gene="At5g10170"
 /evidence=experimental
 /product="putative myo-inositol-1-phosphate synthase"
 /protein_id="AA020204.1"
 /db_xref="GI:20465489"
 /translation="MTESFRVSPHYRGAAIESDYDTTELTVSHSDGASRVV
 RPKSVNFRRTTTPKLGVMVGWGNNGSTLTAGVIANRGI SWATDKYQOANY
 GSLTQASTIRVGSYNGEIEIYAPFKSLIPVNPDLVFGMDISNNLADAMRAKYLD
 IDLOKQRLPYMESMVLPGLIYDPDLFIAANGSRANVIGTKKEQEOIIOIRBEKE
 KSKVDKVVYLVMTANTERSNVYGVNDTENTMILASVDKRAEISPTIYAIAVCVMEGI
 PTNGSPQNTFVPGILDLAIKKNCLIGGDFPSGQTKKSVLVDPLVGAIGKPTSVS
 VNLGNNDGMNLSAPQTFRSKEISNSVDDVANGSLIYEPGHPDHVVVIVKYVCV
 GDSKRAMDEYTSIEIFMGSKNTIVMNTCEDSLAIPILIDVLVLAELTRIDFMSENE
 EKFSHFHPVATILSYLTKAPLVPGTPVNVNALAKORAMLENIVRACVGLAPENNMILE
 YK"
 1534.1564
 /gene="At5g10170"

ORIGIN
 3' UTR

QY 966 GATGACTTCAGAGTGTGACACCA 991
 913 GATGACTTCAGAGTGTGACACCA 938

RESULT 14
LOCUS AX054630 1781 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 1 from Patent WO0073473.
ACCESSION AX054630
VERSION AX054630.1 GI:12228190
KEYWORDS

SOURCE
ORGANISM Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
AUTHORS Georges,F., Hussain,A.A. and Keller,W.A.
TITLE Method for reducing phytate in canola meal using genetic
manipulation involving myo-inositol 1-phosphate synthase gene
JOURNAL Patent: WO 0073473-A 1 07-DEC-2000;
NATIONAL RESEARCH COUNCIL CANADA (CA)
location/Qualifiers
source 1. .1781
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"

ORIGIN
Query Match 1.5%; Score 26; DB 6; Length 1781;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 960 GGTGAGATGACTTCAGAGTGCTCA 985
|||||
Db 963 GGTGAGATGACTTCAGAGTGCTCA 988

RESULT 15
BN066307 1781 bp mRNA linear PLN 02-MAR-2000
LOCUS Brassica napus myo-inositol 1-phosphate synthase mRNA, complete
DEFINITION cds.
ACCESSION U66307
VERSION U66307.1 GI:1513227
KEYWORDS
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 1781)
AUTHORS Hussain,A., Bourgeois,J., Polvi,S., Tsang,E., Keller,W.A. and
Georges,F.
TITLE Cloning of a full length cDNA encoding myo-inositol 1-phosphate
synthase from Brassica napus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1781)
AUTHORS Hussain,A., Bourgeois,J., Polvi,S., Tsang,E., Keller,W.A. and
Georges,F.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1996) National Research Council of Canada, Plant
Biotechnology Institute, 110 Gymnasium Place, Saskatoon,
Saskatchewan S7N 0W9, Canada
location/Qualifiers
source 1. .1781
/organism="Brassica napus"
/mol_type="mRNA"
/db_xref="taxon:3708"
57. .1589
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAB06756.2"
/db_xref="GI:7144546"
/translation="MFIESFKYBSPNVKXTENSHSVYDETTEVVENGAYQWIV
KKRVVYDFKTDIVRKLVGLVGVWGNGSGTILVAGVIANKEGISWATKQVQANYF
GSLTQASIRVGSFNGEEMYPFKSLVPMVNPDDVVGMDISDMVLADAMGRARVLD
IDLOKOLRPYMENTIVPLPGIYDPDFIAAQGRANNVIGTKKEQVDOIIKDMREKKE
KKKVDKVVVMTANTERYGNVIVGLDTEMENLMSVDRDESEISPELTVAIACVLEGI
PRINGSRPTFVPGILDLAIKNNVLIGDDPFKSCOTKMSVLDPLVAGKIGTSTIVS
YVHLGNDDGNTLSAPOTFRSKLSKSNVYDDMVASNGIILFPEGEHPDHVYIKYPIYV
AUSKRAMDEYTSIEFWGKNITVHNATCEDSLAAPIILDVLVAISTRIQKSEKE
GKFHSHFPVATKLSYLTKAPLVPGTPVYNALSKORAMLENTLRACVGLAPENNMILE
YK"

ORIGIN
Query Match 1.5%; Score 26; DB 8; Length 1781;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 960 GGTGAGATGACTTCAGAGTGCTCA 985
|||||
Db 963 GGTGAGATGACTTCAGAGTGCTCA 988

Search completed: June 8, 2005, 11:26:37
Job time : 7685.29 secs


```
Db 1081 AACGTTGATGATATGTCACACAGCAATGCCATCTCTATAGCTGGTGAACTCCA 1140
Qy 1194 GACCATGTTGTTTATTAAATGATGTCCTTACGTAGGGGACAGCAAGAGCCATGAT 1253
Db 1141 GACCATGTTGTTTATTAAATGATGTCCTTACGTAGGGGACAGCAATAGCCATGAT 1200
Qy 1254 GAGTACACTTCAGAGATATTCATGGGTGGAAGAGACCACTGTTTGACACATATCC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAGACCACTGTTTGACACATATCC 1260
Qy 1314 GAGGATTCCTCTTACCTGCTCTTATTTCTTGAATGATGCTCTTCTGAGTCAAG 1373
Db 1261 GAGGATTCCTCTTACCTGCTCTTATTTCTTGAATGATGCTCTTCTGAGTCAAG 1320
Qy 1374 ACTAGATGAGTTTAAAGTGAATAATGAGGAAAAATTCACATCTATCCACCATGTTCT 1433
Db 1321 ACTAGATGAGTTTAAAGTGAATAATGAGGAAAAATTCACATCTATCCACCATGTTCT 1380
Qy 1434 ACCATCTGAGTACTCACTCAACCAAGGCTCTCTGTTCCACCGGTTACACAGTGTGAT 1493
Db 1381 ACCATCTGAGTACTCACTCAACCAAGGCTCTCTGTTCCACCGGTTACACAGTGTGAT 1440
Qy 1494 GCAATTCGAAGCAGGCTGCAATGCTGGAATAATGAGGCTTGTGTTGATTGGCC 1553
Db 1441 GCAATTCGAAGCAGGCTGCAATGCTGGAATAATGAGGCTTGTGTTGATTGGCC 1500
Qy 1554 CCAGAGATPAACATGATTTCTCGATCAACAGTGA 1586
Db 1501 CCAGAGATPAACATGATTTCTCGATCAACAGTGA 1533
```

RESULT 7

ADQ14494

ADQ14494 standard; cDNA; 1533 BP.

XX ADQ14494;

XX 23-SEP-2004 (first entry)

XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.

XX Soybean: myo-inositol 1-phosphate synthase; gene; ss;

XX myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;

XX inorganic phosphate; mutant.

XX Glycine max.

XX Synthetic.

XX OS

XX FH

XX Key

XX CDS

XX 1.1533

XX /tag= a

XX /product= "Mutant soybean myo-inositol 1-phosphate

XX synthase #1"

XX replace (1241,G)

XX mutation

XX /tag= b

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 26-APR-1999; 99US-00293315.

XX 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX (GRAC/) GRACE D J.

XX (STRE/) STREIT L G.

XX XX

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX MPI, 2004-533135/51.

XX F-PSDB; ADQ14495.

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,

XX useful for producing plants with decreased raffinose, stachyose, and

XX phytic acid and increased sucrose, leading to valuable and useful soybean

XX products.

XX Claim 10; SEQ ID NO 5; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-

XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate

XX synthase having a decreased capacity for the synthesis of myo-inositol 1-

XX phosphate. The invention also relates to a chimeric gene operably linked

XX to suitable regulatory sequences, where expression of the chimeric gene

XX results in a decrease in expression of an endogenous or native gene

XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant

XX comprising the chimeric gene, a method of making the soybean plant, a

XX seed of the soybean plant, a soy protein product derived from the

XX processing of soybean seeds, a method of making or producing a soy

XX protein product and a method of using a soybean plant homozygous for at

XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having

XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The

XX nucleic acid fragment and methods are useful for producing plants with

XX decreased raffinose, stachyose and phytic acid content and increased

XX sucrose and inorganic phosphate content, leading to valuable and useful

XX soybean products. This sequence represents cDNA encoding a mutant soybean

XX myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

XX Query Match 84.2%; Score 1482; DB 12; Length 1533;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 54 ATGTTTCATGAGAAATTTTAAAGTTGAGTGTCTTATGTAAGTACACCGAGCTGAGAT 113

XX 1 ATGTTTCATGAGAAATTTTAAAGTTGAGTGTCTTATGTAAGTACACCGAGCTGAGAT 60

XX 114 CAGTCGCTGTAACATGAGAAACCAACCGAATCTGTTACGAGAACGAAATGACATAT 173

XX 61 CAGTCGCTGTAACATGAGAAACCAACCGAATCTGTTACGAGAACGAAATGACATAT 120

XX 174 CAGTGAATGTCMAAACCMAATCTGTCAATGCAATTTAAACCAACATCATCTTCT 233

XX 121 CAGTGAATGTCMAAACCMAATCTGTCAATGCAATTTAAACCAACATCATCTTCT 180

XX 234 AAATTAAGGGTAATGCTGTTGGGTGGGTGGAACCAACGCTCAACCTGACCGGTGT 293

XX 181 AAATTAAGGGTAATGCTGTTGGGTGGGTGGAACCAACGCTCAACCTGACCGGTGT 240

XX 294 GTTATTTGCTAACCGAGAGGCAATTCATGGGCTCAAAAGGACCAAGATTCACAGCAAT 353

XX 241 GTTATTTGCTAACCGAGAGGCAATTCATGGGCTCAAAAGGACCAAGATTCACAGCAAT 300

XX 354 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATCCGAGTTGGTCTCTTCCAGGAGAGAA 413

XX 301 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATCCGAGTTGGTCTCTTCCAGGAGAGAA 360

XX 414 ATCTATGCCCCCAATTAAGAGCTGCTTCCATGTTTAAACCTGACGACATTTGTTGGG 473

XX 361 ATCTATGCCCCCAATTAAGAGCTGCTTCCATGTTTAAACCTGACGACATTTGTTGGG 420

XX 474 GGAATGGAATATGAGCAATGAACTGCTGATGTCATGAGGCGCAGGAGCAAGGTTTGAAC 533

XX 421 GGAATGGAATATGAGCAATGAACTGCTGATGTCATGAGGCGCAGGAGCAAGGTTTGAAC 480

XX 534 ATGATTTTGCAAGAACAGTTGAGGCTTTAATGATGATTCATGCTCCCGGAAATC 593

XX 481 ATGATTTTGCAAGAACAGTTGAGGCTTTAATGATGATTCATGCTCCCGGAAATC 540

QY 594 TATGACCCGGATTTCATTCCTGCCAACAGAGAGACGGTGCACCAACCTCATCAAGGCG 653
DB 541 TATGACCCGGATTTCATTCCTGCCAACAGAGAGACGGTGCACCAACCTCATCAAGGCG 600
QY 654 ACAAAGCAAGACCAAGTTCACCAATCATCAAGAGACATCAAGCGCTTTAAGAACCCACC 713
DB 601 ACAAAGCAAGACCAAGTTCACCAATCATCAAGAGACATCAAGCGCTTTAAGAACCCACC 660
QY 714 AAAGTGAACAAGGTGGTGTACGTGAGCTGCAACACAGAGAGTACAGTAATTTGGTT 773
DB 661 AAAGTGAACAAGGTGGTGTACGTGAGCTGCAACACAGAGAGTACAGTAATTTGGTT 720
QY 774 GTGGGCTTTAATGACCAATGAGAAATCTCTTGCGCTGCTGTGAGACAGAAATAGCGTGA 833
DB 721 GTGGGCTTTAATGACCAATGAGAAATCTCTTGCGCTGCTGTGAGACAGAAATAGCGTGA 780
QY 834 ATTTCTCTTCCACCTTGATGCCATGCTGTGTGTATGGAATAATGCTTCCATTAT 893
DB 781 ATTTCTCTTCCACCTTGATGCCATGCTGTGTGTATGGAATAATGCTTCCATTAT 840
QY 894 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATGATCTTGCCATCGCAGAACACT 953
DB 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATGATCTTGCCATCGCAGAACACT 900
QY 954 TTGATTGTGAGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTGTTGAT 1013
DB 901 TTGATTGTGAGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTGTTGAT 960
QY 1014 TTCTTGTGGGGGCTGTGTATCAAGCCAAATCTATATGATCAAGTACCAATCTGGGAAC 1073
DB 961 TTCTTGTGGGGGCTGTGTATCAAGCCAAATCTATATGATCAAGTACCAATCTGGGAAC 1020
QY 1074 AATGATGTATGATATCTTTCGCGTCCACAAACTTTCCGTTCCAAAGAAATCTCAAGAGC 1133
DB 1021 AATGATGTATGATATCTTTCGCGTCCACAAACTTTCCGTTCCAAAGAAATCTCAAGAGC 1080
QY 1134 AAGTGTGTATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1193
DB 1081 AAGTGTGTATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1140
QY 1194 GACCATGTGTGTATATTAAGTATGCTTACGTAAGGGGACAGCAAGAGCCATGAT 1253
DB 1141 GACCATGTGTGTATATTAAGTATGCTTACGTAAGGGGACAGCAATGAGCCATGAT 1200
QY 1254 GAGTACACTTCAAGATATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
DB 1201 GAGTACACTTCAAGATATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1314 GAGGATTCCTCTTACCTGCTCTATATCTTGAATGATGATGATGATGATGATGATGAT 1373
DB 1261 GAGGATTCCTCTTACCTGCTCTATATCTTGAATGATGATGATGATGATGATGATGAT 1320
QY 1374 ACTAGAAATGAGATTAAGTAAAGTAAATGAGGAAATTCACATTCATCCAGTTGCT 1433
DB 1321 ACTAGAAATGAGATTAAGTAAAGTAAATGAGGAAATTCACATTCATCCAGTTGCT 1380
QY 1434 ACCATCTCAGTACCTCAGCAAGAGCTCTCTGTTTCCACCGGGTACACAGTGTGAAT 1493
DB 1381 ACCATCTCAGTACCTCAGCAAGAGCTCTCTGTTTCCACCGGGTACACAGTGTGAAT 1440
QY 1494 GCATTGTCAAAAGCAGGTGCAATGCTGGAATAATGAGGGCTGTGTGATTGGCC 1553
DB 1441 GCATTGTCAAAAGCAGGTGCAATGCTGGAATAATGAGGGCTGTGTGATTGGCC 1500
QY 1554 CCAGAGAAATTAACATGATTTCTCAGTACCAAGTGA 1586
DB 1501 CCAGAGAAATTAACATGATTTCTCAGTACCAAGTGA 1533

RESULT 8
ADS81997
ID ADS81997 standard; cDNA; 1533 BP.
XX

AC ADS81997;
XX
XX 18-NOV-2004 (first entry)
DT
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
XX Glycine max; line LR33.
OS
XX
XX Synthetic.
FH
FH Key Location/Qualifiers
FT CDS 1..1533
FT /*tag= a
FT /product= "myo-inositol 1-phosphate synthase"
FT replace(1188,G)
FT /*tag= b
FT mutation
PN
XX US2003074685-A1.
XX
XX 17-APR-2003.
PD
XX
XX 11-MAR-2002; 2002US-00025003.
PF
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
PI
XX
XX WPI: 2004-639957/62.
DR P-PsDB; ADS81998.
DR
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
PT
XX
PS Claim 8; SEQ ID NO 5; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g), provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a mutant myo-inositol 1-phosphate synthase.
XX

Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 84.2%; Score 1482; DB 13; Length 1533;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 54 ATGTCATGAGAAATTTTAAAGTTGAGTGCCTTAATGTGAAGTACCCGAGCTGAGATT 113
DB 1 ATGTCATGAGAAATTTTAAAGTTGAGTGCCTTAATGTGAAGTACCCGAGCTGAGATT 60
QY 114 CAGTCCGTGACAACTACGAAACACCGAATTGTTCAAGAGAACGAAATGACCTAT 173
DB 61 CAGTCGTGACAACTACGAAACACCGAATTGTTCAAGAGAACGAAATGACCTAT 120
QY 174 CAGTGAATGTCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAATCATGTTCT 233
DB 121 CAGTGAATGTCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAATCATGTTCT 180
QY 234 AAATTAGGGGTATGCTGTGGGGTGGGAGAAACAACGCTCAACCCCTCAACGGTGT 293
DB 181 AAATTAGGGGTATGCTGTGGGGTGGGAGAAACAACGCTCAACCCCTCAACGGTGT 240
QY 294 GTTATGCTAACCGAGAGGCTATTCATGAGGCTACAAAGGACAAAGATTCAACAGCAAT 353
DB 241 GTTATGCTAACCGAGAGGCTATTCATGAGGCTACAAAGGACAAAGATTCAACAGCAAT 300
QY 354 TACTTTGGCTCCCTCAACCCAGGCTCACTATCCGAGTTGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAGGCTCACTATCCGAGTTGGTCTTCCAGGAGAGAA 360
QY 414 ATCTATGCCCATTCAGAGGCTGCTTCCAAATGTTTAAACCTGACGACATGTTGTTGG 473
DB 361 ATCTATGCCCATTCAGAGGCTGCTTCCAAATGTTTAAACCTGACGACATGTTGTTGG 420
QY 474 GGAATGGATATACGAAACATGAACCTGTGATGCAATGGCCAGGCAAAAGTGTGAC 533
DB 421 GGAATGGATATACGAAACATGAACCTGTGATGCAATGGCCAGGCAAAAGTGTGAC 480
QY 534 ATCGATTTCAGAGGATTTAGGCTTTACATGAAATCCATGCTTCACTCCCGGAATC 593
DB 481 ATCGATTTCAGAGGATTTAGGCTTTACATGAAATCCATGCTTCACTCCCGGAATC 540
QY 594 TATGACCCGGATTTCAATGCTGCGCAACAGAGGACCTGCCAAACACGTCATCAAGGCG 653
DB 541 TATGACCCGGATTTCAATGCTGCGCAACAGAGGACCTGCCAAACACGTCATCAAGGCG 600
QY 654 ACAAAAGCAAGAGCAATTCATCAAAAGATCATCAAGGCTTTTAAAGAGGACAC 713
DB 601 ACAAAAGCAAGAGCAATTCATCAAAAGATCATCAAGGCTTTTAAAGAGGACAC 660
QY 714 AAAGTGACAGGTGTTGTAATCTGTGACTGCCAACAGAGAGTACAGTAATTTGGTT 773
DB 661 AAAGTGACAGGTGTTGTAATCTGTGACTGCCAACAGAGAGTACAGTAATTTGGTT 720
QY 774 GTGGGCTTAAAGACCAATGAGATCTCTTGGCTGCTGTGACAGAAATAGGCTGAG 833
DB 721 GTGGGCTTAAAGACCAATGAGATCTCTTGGCTGCTGTGACAGAAATAGGCTGAG 780
QY 834 ATTTCTCTTCCACCTTGTATGCAATGCTGTGATGGAATAATGTTCTTTCAATTAAT 893
DB 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
QY 894 GGAAGCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAAACACT 953
DB 841 GGAAGCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAAACACT 900
QY 954 TTGATTTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGAT 1013
DB 901 TTGATTTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGAT 960
QY 1014 TTCTTTGAGGAGCTGTGATCAAGGCAATCTATAGTACAGTTACCAACCATCTGGGAAC 1073
DB 961 TTCTTTGAGGAGCTGTGATCAAGGCAATCTATAGTACAGTTACCAACCATCTGGGAAC 1020

```

```

QY 1074 AATGATGATGATATCTTTGGGCTCCACAAACTTTCCGTTCCAGGAATCTTCAAGAGC 1133
DB 1021 AATGATGATGATATCTTTGGGCTCCACAAACTTTCCGTTCCAGGAATCTTCAAGAGC 1080
QY 1134 AACGTTGATGATATGATGCAACAGCAATGCGATCTCTATAGAGCTGTGAACATCCA 1193
DB 1081 AACGTTGATGATATGATGCAACAGCAATGCGATCTCTATAGAGCTGTGAACATCCA 1140
QY 1194 GACCATGTTGTTTATTAATATGTCCTTACGTAAGGAGACAGAAAGAGCATGAT 1253
DB 1141 GACCATGTTGTTTATTAATATGTCCTTACGTAAGGAGACAGCAATGAGCATGAT 1200
QY 1254 GAGTACACTTCAGAGATATTCATGAGTGAAGAGCAACATGTTTTCACAAACATGC 1313
DB 1201 GAGTACACTTCAGAGATATTCATGAGTGAAGAGCAACATGTTTTCACAAACATGC 1260
QY 1314 GAGGATTCCTCTAGCTGCTCTATTAATCTTGAATGTCCTTCTGCTGAGCTCAGC 1373
DB 1261 GAGGATTCCTCTAGCTGCTCTATTAATCTTGAATGTCCTTCTGCTGAGCTCAGC 1320
QY 1374 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTCACATTCACACCAATGCT 1433
DB 1321 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTCACATTCACACCAATGCT 1380
QY 1434 ACCATCTCAGTACTCTCAACAGGCTCTCTGTGTTCCACCGGATACCAAGTGTGAAT 1493
DB 1381 ACCATCTCAGTACTCTCAACAGGCTCTCTGTGTTCCACCGGATACCAAGTGTGAAT 1440
QY 1494 GCATTTGCAAGCAGGTCATATGCTGGAATAATATAGGCTGTGTGGAATGGGC 1553
DB 1441 GCATTTGCAAGCAGGTCATATGCTGGAATAATATAGGCTGTGTGGAATGGGC 1500
QY 1554 CCAGAGAAATACATGATTTCTCGAGTACAAAGTGA 1586
DB 1501 CCAGAGAAATACATGATTTCTCGAGTACAAAGTGA 1533

```

RESULT 9
 ADQ14504
 ID ADQ14504 standard; cDNA; 1533 BP.
 XX
 AC ADQ14504;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 KW inorganic phosphate.
 XX
 OS Glycine max.
 XX
 PH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "Wild type soybean myo-inositol 1-phosphate
 FT synthase #2"
 FT
 PN US2004128713-A1.
 XX
 XX 01-JUL-2004.
 PD
 XX
 PF 21-NOV-2003; 2003US-00718952.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98MO-US006822.
 PR 26-APR-1999; 99US-0029315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI; 2004-53335/51.
DR P-PSDB; ADQ14505.
XX
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX
PS Claim 4; SEQ ID NO 15; 48bp; English.
XX
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant,
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 8.3%; Score 146; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 2.3e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 ATGAAAATGTTCTTCTTCAATTAATGAAGCCCTCAGAACCTTTGTACAGGGCTGATT 929
DB 817 ATGAAAATGTTCTTCTTCAATTAATGAAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
QY 930 GATCTTGCATCGCGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGCC 989
DB 877 GATCTTGCATCGCGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGCC 936
QY 990 AAAATGAATCTGTGTGTGTTGATT 1015
DB 937 AAAATGAATCTGTGTGTGTTGATT 962

RESULT 10
ADQ14500 standard; cDNA; 1533 BP.
XX
XX ADQ14500;
XX 23-SEP-2004 (first entry)
XX
XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.
DE
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
XX Glycine max.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT CDS 1..1533
FT /*tag= a

FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #3"
XX
XX
PN US2004128713-A1.
XX
XX
PD 01-JUL-2004.
XX
XX
PF 21-NOV-2003; 2003US-00718952.
XX
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006832.
PR 26-APR-1999; 99US-00289315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI; 2004-53335/51.
DR P-PSDB; ADQ14501.
XX
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX
PS Claim 10; SEQ ID NO 11; 48bp; English.
XX
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 8.3%; Score 146; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 2.3e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 ATGAAAATGTTCTTCTTCAATTAATGAAGCCCTCAGAACCTTTGTACAGGGCTGATT 929
DB 817 ATGAAAATGTTCTTCTTCAATTAATGAAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
QY 930 GATCTTGCATCGCGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGCC 989
DB 877 GATCTTGCATCGCGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGCC 936
QY 990 AAAATGAATCTGTGTGTGTTGATT 1015
DB 937 AAAATGAATCTGTGTGTGTTGATT 962

RESULT 11
ADQ14502 standard; cDNA; 1533 BP.
XX

AC ADQ14502;
XX
XX 23-SEP-2004 (first entry)
XX
XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
DE
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.
XX
XX Glycine max.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #4"
XX
XX US2004128713-A1.
XX
XX 01-UTL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX PF 07-APR-1998; 98WO-US006822.
XX PR 26-APR-1999; 99US-0029315.
XX PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-53135/51.
XX P-PSDB; ADQ14503.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Example 8; SEQ ID NO 13; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents cDNA encoding a mutant soybean
XX myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 8.3%; Score 146; DB 12; Length 1533;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-60;
XX Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 870 ATGGAATATGTTCTTCAATATGGAACCCCTCAGAACACTTTGTACAGGCTGATT 929

DB 817 ATGGAATATGTTCTTCAATATGGAACCCCTCAGAACACTTTGTACAGGCTGATT 876
QY 930 GATCTTGCATGCGGAGAACACTTTGATTGTGAGATGACTTCAAGAGTGTACAGC 969
DB 877 GATCTTGCATGCGGAGAACACTTTGATTGTGAGATGACTTCAAGAGTGTACAGC 936
QY 990 AAAATGAATCTGTGTGTTGTTGATT 1015
DB 937 AAAATGAATCTGTGTGTTGTTGATT 962
RESULT 12
ADSS82005
ID ADSS82005 standard; cDNA; 1533 BP.
XX
XX ADSS82005;
XX
XX 18-NOV-2004 (first entry)
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; cultivar Wye.
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX /tag= a
XX /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
XX P-PSDB; ADSS82006.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 2; SEQ ID NO 15; 34bp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis for myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived

CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 8.3%; Score 146; DB 13; Length 1533;

Best Local Similarity 100.0%; Pred. No. 2.3e-60; Mismatches 0; Gaps 0;

Matches 146; Conservative 0; Indels 0; Gaps 0;

QY 870 ATGAAAATGTTCTTTCATTATGGAAGCCCTCAGAACCTTTGTACAGGGCTGATT 929

DB 817 ATGAAAATGTTCTTTCATTATGGAAGCCCTCAGAACCTTTGTACAGGGCTGATT 876

QY 930 GATCTTGCATGCGAGGAACAACCTTGTGATGATGACTTCAAGATGCTCAGACC 989

DB 877 GATCTTGCATGCGAGGAACAACCTTGTGATGATGACTTCAAGATGCTCAGACC 936

QY 990 AAAATGAATCTGTGTGTTGATT 1015

DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 13

AD882003 standard; cDNA; 1533 BP.

AC AD882003;

DT 18-NOV-2004 (first entry)

DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.

KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;

KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

OS Glycine max; line 29018BP03.

Key Location/Qualifiers

FT CDS 1..1533

FT /product= a

FT /product= "myo-inositol 1-phosphate synthase"

PN US2003074685-A1.

PD 17-APR-2003.

PF 11-MAR-2002; 2002US-00025003.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PI Hitz WD, Sebastian SA;

XX WPI; 2004-63957/62.

XX P-PSDB; AD882004.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.

PS Example 8; SEQ ID NO 13; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis for myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment), operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 17 micromol/9, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/9, and a seed sucrose content of
XX greater than 200 micromol/9, provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived
XX from seeds of a soybean plant (homozygous for one or more gene encoding a
XX mutant myo-inositol 1-phosphate synthase having decreased capacity for
XX the synthesis of myo-inositol 1-phosphate, where the gene confers a
XX heritable phenotype as mentioned above), and making or producing a
XX soybean protein product derived from seeds of a soybean plant with a
XX heritable phenotype as mentioned above. The nucleic acid is useful for
XX altering raffinose saccharide, sucrose, phytic acid and inorganic
XX phosphate content of soybean seeds thus leading to valuable and useful
XX soybean products, since the presence of high concentration of raffinose
XX oligosaccharides in soy plants (and other legumes) can lead to flatulence
XX when consumed by humans. The present sequence encodes a wild-type myo-
XX inositol 1-phosphate synthase.

SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 8.3%; Score 146; DB 13; Length 1533;

Best Local Similarity 100.0%; Pred. No. 2.3e-60; Mismatches 0; Gaps 0;

Matches 146; Conservative 0; Indels 0; Gaps 0;

QY 870 ATGAAAATGTTCTTTCATTATGGAAGCCCTCAGAACCTTTGTACAGGGCTGATT 929

DB 817 ATGAAAATGTTCTTTCATTATGGAAGCCCTCAGAACCTTTGTACAGGGCTGATT 876

QY 930 GATCTTGCATGCGAGGAACAACCTTGTGATGATGACTTCAAGATGCTCAGACC 989

DB 877 GATCTTGCATGCGAGGAACAACCTTGTGATGATGACTTCAAGATGCTCAGACC 936

QY 990 AAAATGAATCTGTGTGTTGATT 1015

DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 14

AD882001 standard; cDNA; 1533 BP.

AC AD882001;

DT 18-NOV-2004 (first entry)

DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.

KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;

KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.

OS Glycine max; line 29010CP01.

OS Synthetic.

XX Key

XX Location/Qualifiers

XX FT CDS 1..1533

XX /product= a

XX /product= "myo-inositol 1-phosphate synthase"

XX mutation


```

FT      /+tag= b
XX      US2003074685-A1.
XX      17-APR-2003.
XX      11-MAR-2002; 2002US-00025003.
XX      08-APR-1997; 97US-00835751.
XX      07-APR-1998; 98WO-US006822.
XX      (HITZ/) HITZ W D.
XX      (SEBA/) SEBASTIAN S A.
XX      Hitz WD, Sebastian SA;
XX      MPI; 2004-639957/62.
XX      P-PSDB; ADS82002.
XX      Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT      phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT      phytic acid and inorganic phosphate content of soybean seeds.
XX      Claim 8; SEQ ID NO 11; 34pp; English.
XX      The invention relates to an isolated nucleic acid fragment encoding a
CC      soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC      phosphate synthase having decreasing capacity for the synthesis for myo-
CC      inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC      nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC      complement, subfragment or the complement of the subfragment, operably
CC      linked to suitable regulatory sequences, where expression of the chimeric
CC      gene results in a decrease in expression of an endogenous or native gene
CC      encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC      comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC      acid content of less than 17 micromol/g, a seed content of raffinose plus
CC      stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC      greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC      from the plant, making a soybean plant with the heritable phenotype
CC      (comprising crossing LR33 or the plant comprising the chimeric gene with
CC      an elite soybean plant and selecting a progeny plant of the cross of
CC      crossing step that has a heritable phenotype as mentioned above), seeds
CC      of soybean plant made by the above method, a soy protein product derived
CC      from seeds of a soybean plant (homozygous for one or more gene encoding a
CC      mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC      the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC      heritable phenotype as mentioned above), and making or producing a
CC      soybean protein product derived from seeds of a soybean plant with
CC      heritable phenotype as mentioned above. The nucleic acid is useful for
CC      altering raffinose saccharide, sucrose, phytic acid and inorganic
CC      phosphate content of soybean seeds thus leading to valuable and useful
CC      soybean products, since the presence of high concentration of raffinose
CC      oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC      when consumed by humans. The present sequence encodes a mutant myo-
CC      inositol 1-phosphate synthase.
XX      Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;
SQ
Query Match      8.3%; Score 146; DB 13; Length 1533;
Best Local Similarity 100.0%; Pred. No. 2.3e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      870 ATGGAATAATGTTCTTTATTAAATGGAAGCCTCAGAAACACTTTTACCAAGGCTGATT 929
XX      |||||||
XX      DB      817 ATGGAATAATGTTCTTTATTAAATGGAAGCCTCAGAAACACTTTTACCAAGGCTGATT 876
XX      |||||||
XX      QY      930 GATCTGGCATGCGAGAGAACTTGGATGTGGAGAGACTTCAAGAGTGTTCAGACC 989
XX      |||||||
XX      DB      877 GATCTGGCATGCGAGAGAACTTGGATGTGGAGAGACTTCAAGAGTGTTCAGACC 936
XX      |||||||
XX      QY      990 AAAATGAATCTGTGTGGTTGATT 1015
XX      |||||||
XX      DB      937 AAAATGAATCTGTGTGGTTGATT 962

```

```

RESULT 15
ADP93370
ID      ADP93370 standard; cDNA; 377 BP.
XX
XX      AC      ADP93370;
XX      DT      09-SEP-2004 (first entry)
XX      XX      Cotton expressed sequence tag, EST, #2381.
XX      DE      Cotton; ss; EST; expressed sequence tag; plant; plant protection;
XX      KM      plant improvement; marker-assisted breeding.
XX      OS      Gossypium hirsutum; variety Nucleon33B.
XX      PN      US2004123338-A1.
XX      PD      24-UTN-2004.
XX      PF      08-DEC-2000; 2000US-00732627.
XX      PR      10-DEC-1999; 99US-0170255P.
XX      PA      (FINC/) FINCHER K L.
XX      PI      Fincher KL;
XX      DR      MPI; 2004-479807/45.
XX      PT      New substantially purified nucleic acid molecule that encodes a cotton
XX      protein or its fragment, useful as molecular tool for the targeting and
XX      isolation of novel genes for plant protection and improvement.
XX      PS      Claim 1; SEQ ID NO 2381; 30pp; English.
XX      The invention relates to a substantially purified nucleic acid molecule
CC      that encodes a cotton protein or its fragment comprising an EST
CC      (expressed sequence tag) appearing as ADP90990-ADP95919. Also included
CC      are a substantially purified cotton protein or its fragment encoded by a
CC      nucleic acid molecule above and a transformed plant (having a nucleic
CC      acid molecule which comprises: an exogenous promoter region which
CC      functions in a plant cell to cause the production of a mRNA molecule; a
CC      structural nucleic acid molecule comprising one of the ESTs or their
CC      complements; a 3' non-translated sequence that functions in the plant
CC      cell to cause termination of transcription and addition of polyadenylated
CC      ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
CC      molecular tool for the targeting and isolation of novel genes for plant
CC      protection and improvement. The ESTs are useful for developing new
CC      strategies for understanding critical plant developmental and metabolic
CC      pathways, for isolating genes and promoters, for identifying and mapping
CC      the genes involved in developmental and metabolic pathways, and for
CC      determining gene function. The cotton nucleic acid molecules are useful
CC      as molecular tags to isolate genetic regions, isolate genes, map genes,
CC      and determine gene function. The nucleic acid molecules are useful for
CC      determining if genes are members of a particular gene family and for use
CC      in marker-assisted breeding programs. The present sequence is one of the
CC      4930 cotton ESTs of the invention. NOTE: The sequences are not displayed
CC      in the specification but are available in electronic format from the
XX      USPTO at seqdata.uspto.gov/sequence.html?DocID=20040123338.
XX      SQ      Sequence 377 BP; 100 A; 87 C; 90 G; 100 T; 0 U; 0 Other;
XX
Query Match      1.6%; Score 29; DB 12; Length 377;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      1233 GACAGCAAGAGAGCCATGATGATGATCAC 1261
XX      |||||||
XX      DB      161 GACAGCAAGAGAGCCATGATGATGATCAC 189

```


Search completed: June 8, 2005, 07:26:59
Job time : 970.851 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:49:46 ; Search time 5995.12 Seconds

(without alignment)
11174.613 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttctctctcttctgt.....aatgtagtataatttctgt 1760

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	750	42.6	813	7	CK768601 Gm-r1030-
2	552	31.9	686	2	AM348857 Gm-r10010A
3	533	30.3	533	4	BM523576 bam85904.
4	527	29.9	672	2	AM508269 s151n10.Y
5	517	29.4	622	2	AM397496 sg79e12.Y
6	508	28.9	619	2	BE313163 sg98f01.Y
7	505	28.7	575	2	AM397453 sg79d01.Y
8	501	28.5	648	1	AT794750 sg58f07.Y
9	475	27.0	539	7	AM317927 sg58f08.Y
10	469	26.6	459	7	CN472440 Gm-r1030-
11	463	26.5	478	7	CK769092 Gm-r1030-
12	443	25.2	443	5	BO612082 bap77b07.
13	400	22.7	420	4	BI347339 F165E19.C
14	391	22.2	435	2	BE609839 sg45a10.Y
15	383	21.8	487	2	AM472457 sg125d05.Y
16	375	21.3	496	2	AM509380 s122d05.Y
17	363	20.6	451	4	BG726672 ssa31b08.
18	359	20.4	576	1	AI496426 sb04e09.Y
19	315	17.9	316	2	AM398011 sg71b06.Y
20	299	17.0	308	2	AM705757 sg51e05.Y
21	290	16.5	290	2	AM460108 s110d02.Y
22	279	15.9	408	5	BU763677 ssa84d09.
23	276	15.7	661	2	BE191464 sn76h10.Y
24	246	14.0	552	4	BM308082 ssa41f07.

25	243	13.8	396	2	AM472088	AM472088 s119g09.Y
26	205	11.6	208	2	AM569076	AM569076 s163b05.Y
27	201	11.4	201	4	BG404951	BG404951 s8c4d03.Y
28	172	9.8	271	2	AM568795	AM568795 s161b10.Y
29	152	8.6	152	4	BG237745	BG237745 s8b17a01.
30	152	8.6	152	4	BG651312	BG651312 s8d83a07.
31	146	8.3	304	2	BP425513	BP425513 s85f04.Y
32	146	8.3	449	2	AM203517	AM203517 s8f35b05.Y
33	146	8.3	516	4	BG652636	BG652636 s8d69g02.
34	146	8.3	540	5	BM887128	BM887128 s8m34e09.
35	146	8.3	552	5	BM955039	BM955039 s8m76c03.
36	146	8.3	634	2	BE331050	BE331050 s8c93e10.Y
37	145	8.2	670	4	BG044525	BG044525 s8a29e07.
38	143	8.1	563	2	AM472315	AM472315 s1233f04.Y
39	138	7.8	158	2	BE603020	BE603020 29 GmaxSC
40	135	7.7	388	2	AM432728	AM432728 s8h5f08.Y
41	135	7.7	411	2	AM099866	AM099866 s8l17e12.Y
42	135	7.7	420	4	BM092617	BM092617 s8h16d10.
43	135	7.7	456	2	AM620996	AM620996 s150e08.Y
44	135	7.7	457	2	AM310369	AM310369 s8f35b05.X
45	135	7.7	467	1	AT442850	AT442850 s8a27f05.X

ALIGNMENTS

RESULT 1
CK768601
LOCUS
DEFINITION
Gm-r1030-621 Gm-r1030 Glycine max cDNA clone Gm-r1030-621 5', mRNA
sequence.
CK768601
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CK768601.1 GI:42722702
EST.
Glycine max (soybean)
Glycine max

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 813)
Periapuram,C.C., Li,L., Wurttele,E.S., Westgate,M.E. and Nikolau,B.J.
Molecular determination of soybean composition
Unpublished (2004)
Contact: Basil J. Nikolau and Cyril C. Periapuram
Department of Biochemistry, Biophysics and Molecular Biology
Iowa State University
2210 Molecular Biology Building, Ames, IA 50011, USA
Tel: 515 294 9423
Fax: 515 294 0453
Email: dimmas@iastate.edu

Individual base call and confidence value were assigned using the Phred software (http://www.phrap.org/). Overall sequence quality assessment and vector trimming were conducted using the Lucy software (http://www.tigr.org/software/).
This clone was originally generated by the Public Soybean EST Project (http://129.186.26.94/soybeanest.html)/Shoemaker, R (rcshoe@iastate.edu).
This clone is available through: Biogenetic Services, 801 32nd Ave, Brookings, SD 57006. For further information call 605-697-8500 or contact info@biogeneticservices.com
Seg primer: T7-1 (5' AAT ACG ACT CAC TAT AG 3').

FEATURES

source
1..813
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1030-621"
/lab_host="DH10B"
/clone_1ib="Gm-r1030"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This cDNA library was constructed from mRNA isolated from

immature cotyledons of greenhouse grown plants (individual seed fresh weight of 100-300mg). The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restriction site. SalI linker adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note that Gm-r1021 is a re-rack of Gm-cl007."

ORIGIN

Query Match 42.6%; Score 750; DB 7; Length 813;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

487 GCAACATGAACCTGGCTGATGCGATGCGCAGGCAAGGCTTTGACATGATTTGACAGA 546
13 GCAACATGAACCTGGCTGATGCGATGCGCAGGCAAGGCTTTGACATGATTTGACAGA 72
547 AGCAGTTGAGGCTTACATGATGATTCATCTCCCGAATCTATGACCCGGAAT 606
73 AGCAGTTGAGGCTTACATGATGATTCATCTCCCGAATCTATGACCCGGAAT 132
607 TCATTTGCTGCGCAACGAGAGGCGCGCAACAGCTCATCAAGGCGCAACAAAGAGC 666
133 TCATTTGCTGCGCAACGAGAGGCGCGCAACAGCTCATCAAGGCGCAACAAAGAGC 192
667 AAGTTCAACAAATCATCAAGACATCAAGAGCGCTTTAAGAGACCAACAAAGTGAACAAG 726
193 AAGTTCAACAAATCATCAAGACATCAAGAGCGCTTTAAGAGACCAACAAAGTGAACAAG 252
727 TGGTGTACTGTGAGCTGCCAACACAGAGAGGTACAGTAATTTGTTGGGCGCTTAATG 786
253 TGGTGTACTGTGAGCTGCCAACACAGAGAGGTACAGTAATTTGTTGGGCGCTTAATG 312
787 AACACATGAGAAATCTTTGGCTGCTGTGAGACAGAAATAGAGCTGATTTCTCCCTCCA 846
313 AACACATGAGAAATCTTTGGCTGCTGTGAGACAGAAATAGAGCTGATTTCTCCCTCCA 372
847 CCTTGATGCCATTTGCTGTGATGAGAAATGTTCTCTTCAATTAATGAGAGCCCTCAGA 906
373 CCTTGATGCCATTTGCTGTGATGAGAAATGTTCTCTTCAATTAATGAGAGCCCTCAGA 432
907 AACATTTGTACAGGCGCTGATTTGATCTTGCATGCGAGAAACATTTGATTTGGTGA 966
433 AACATTTGTACAGGCGCTGATTTGATCTTGCATGCGAGAAACATTTGATTTGGTGA 492
967 ATGACTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGTTGTTTCTTTGGGGGG 1026
493 ATGACTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGTTGTTTCTTTGGGGGG 552
1027 CTGATCAAGGCAACATCTATAGTCACTTAACACATCTGGGAAACAAATGATGATGA 1086
553 CTGATCAAGGCAACATCTATAGTCACTTAACACATCTGGGAAACAAATGATGATGA 612
1087 ATCTTTGGGCTCCAAACATTTCCGTTCCAGAGAAATCTCCAAAGCAACGTTTGTGATG 1146
613 ATCTTTGGGCTCCAAACATTTCCGTTCCAGAGAAATCTCCAAAGCAACGTTTGTGATG 672
1147 AATAGTCAACAGCAATGCGATCTCTATGAGCGTGTGAACATCCAGACATGTTGTTG 1206
673 AATAGTCAACAGCAATGCGATCTCTATGAGCGTGTGAACATCCAGACATGTTGTTG 732
1207 TTATTAGATGTCCTTAGTAGGAGAGCAAGAGACCATGATGATGATCACTTCA 1266
733 TTATTAGATGTCCTTAGTAGGAGAGCAAGAGACCATGATGATGATCACTTCA 792
1267 AGATATTCATGGGTGAAAGA 1287

Db 793 AGATATTCATGGGTGAAAGA 813

RESULT 2
LOCUS AM348857/c 686 bp mRNA linear EST 04-OCT-2000
DEFINITION GM210010A10E12 Gm-r1021 Glycine max cDNA clone Gm-r1021-3671 3',
MRNA sequence.
AM348857
AM348857.1 GI:6846567
EST.
Glycine max (soybean)
SOURCE
ORGANISM
Glycine max
Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseolae;
Glycine.
1 (bases 1 to 686)
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
Erpelding, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.
A Functional Genomics Program for Soybean (NSF 9872565)
Unpublished (1999)
Other ESTs: AL496426
Contact: Vodkin, L.O., PI, A Functional Genomics Program for
Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional
Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu

REFERENCES

AUTHORS

TITLE

JOURNAL

COMMENT

This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: cloness@genomesystems.com or info@genome
systems.com web site: www.genomesystems.com
Seq primer: 5'-TTTCTTTTCTTTTCTTTT(A/C/G)-3'.

FEATURES

source

1..686
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="Gm-r1021-3671"
/tissue_type="root"
/lab_host="XLL10-Gold"
/clone_lib="Gm-r1021"
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
XhoI; Library Gm-r1021 is a sequence-driven, re-racked set
of the original library Gm-cl004 which was prepared from
root cDNA. The mRNA was isolated from entire roots of 8
day old 'Williams' seedlings which were propagated on
paper towels with distilled water. StrataGene 8 cDNA
Synthesis Kit (catalog #200401) was used to synthesize
the cDNA. The Gm-cl004 library was constructed by Dr.
Paul Keim & Virginia H. Coryell, Department of Biology,
Box5640, Northern Arizona University, Flagstaff, AZ
86011, email: paul.keim@uau.edu, virginia.coryell@uau.edu.
The contig analysis to select unique genes was performed
by the laboratory of Ernest Retzel, Computational Biology
Centers, University of Minnesota,
http://www.cb.c.umn.edu/ResearchProjects/Soybean/index.html
Retzel was performed by Genome Systems, St. Louis,
http://www.genomesystems.com, and sequencing by the Keck
Center for Comparative and Functional Genomics,
University of Illinois,
http://www.life.uiuc.edu/biotech/keck.html."

ORIGIN

Query Match 31.9%; Score 562; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 5.2e-305; Indels 0; Gaps 0;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

```

FEATURES
source
1. .672
Location/Qualifiers
Email: est@watson.wastl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1864 Std Error: 0.00
High quality sequence stop: 431.

```

```

/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-r1030-2492"
/lab_host="DH10B"
/clone_id="Gm-r1030"
/notes="vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants
(individual seed fresh weight of 100-300mg). The library
was prepared using the Life Technologies pSP65script cDNA
library construction kit. Complementary DNA was
synthesized from mRNA using a poly(dI) sequence with a
NotI restriction site. SalI linkers adapters were ligated
to the blunt-ended cDNA fragments followed by NotI
digestion. The cDNA fragments were directionally cloned
into the NotI-SalI restriction site of the pSPORT1
vector. The ligated cDNA fragments were transformed into
E. coli Electromax DH10B host cells. This library was
constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note
that Gm-r1030 is a re-rack of Gm-cl007."

```

[illegible]

OY	138	ACCGAATCTGTTCACGGAACAGGAATGGCACTATCAGTGGATGTGCAACCAGAAATCT	197
Db	121	ACCGAATCTGTTCACGGAACAGGAATGGCACTATCAGTGGATGTGCAACCAGAAATCT	180
OY	198	GTCAATAACGAATTTTAAAACCCAATCATGTTCTTAAATTAGGGGTATGCTTGTGGGT	257
Db	181	GTCAATAACGAATTTTAAAACCCAATCATGTTCTTAAATTAGGGGTATGCTTGTGGGT	240
OY	258	TGGGGTGGAAACAACGGCTCAACCTTCACGGTGGTGTATTTGGTAAACCGAGAGGCATTT	317
Db	241	TGGGGTGGAAACAACGGCTCAACCTTCACGGTGGTGTATTTGGTAAACCGAGAGGCATTT	300
OY	318	TCATGGGGTCAAAAGACAAGATTCAACAACCAATTACTTTGGCTCCCTCAACCAAGCC	377
Db	301	TCATGGGGTCAAAAGACAAGATTCAACAACCAATTACTTTGGCTCCCTCAACCAAGCC	360
OY	378	TCAGCTATCCGAGTTGGGTCTTTCACAGGAGAGGAAATCTATGCCCCATTCAAGAGCCTG	437
Db	361	TCAGCTATCCGAGTTGGGTCTTTCACAGGAGAGGAAATCTATGCCCCATTCAAGAGCCTG	420
OY	438	CTTCCAAATGGTTAAACCTTGACGACATTTGTGTTGGGGGATGGGATATTAGCAACATGAAC	497
Db	421	CTTCCAAATGGTTAAACCTTGACGACATTTGTGTTGGGGGATGGGATATTAGCAACATGAAC	480
OY	498	CTGGCTGATGCGATGCGACAGGCAAAAGTGTTTGACATCGATTTTGGCA	544
Db	481	CTGGCTGATGCGATGCGACAGGCAAAAGTGTTTGACATCGATTTTGGCA	527

RESULT 5	
LOCUS	AM397496
DEFINITION	622 bp mRNA linear EST 14-JUL-2004
ACCESSION	AM397496
VERSION	1
KEYWORDS	sgf99e12.y1 Gm-cl007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: P42802
SOURCE	Gm-cl007-2663 5' similar to SW:INOL1 CTPA P42802
ORGANISM	MVO-INOSITOL-1-PHOSPHATE SYNTHASE ;, mRNA sequence.
	AM397496
	AM397496.1 GI:6915366
	EST.
	Glycine max (soybean)
	Glycine max

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 622)
Shoemaker, R., Keim, P., Vodkin, L., Expediting, J., Corryell, V.,
Khanna, A., Bolla, B., Marras, M., Hillier, L., Kueaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gidbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gidco
High quality sequence stop: 434.
Location/Qualifiers
1. .622

```

FEATURES
    source
        location/Qualifiers
            1..622
                /organism="Glycine max"
                /mol_type="mRNA"
                /cultivar="Williams"
                /db_xref="taxon:3847"

```

/clone="GENOME SYSTEMS CLONE ID: Gm-c1007-2663"
/lab_host="DH10B"
/clone_lib="Gm-c1007"
/note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants (individual
seed fresh weight of 100-300mg). The library was prepared
using the Life Technologies pSuperscript cDNA library
construction kit. Complementary DNA was synthesized from
mRNA using a poly (dT) sequence with a NotI restriction
site. SalI linkers adapters were ligated to the
blunt-ended cDNA fragments followed by NotI digestion. The
cDNA fragments were directionally cloned into the
NotI-SalI restriction site of the pSPORT1 vector. The
ligated cDNA fragments were transformed into E. coli
Electromax DH10B host cells. This library was constructed
by Dr. Lila Vodka and Dr. Anu Khanna."

ORIGIN

Query Match 29.4%; Score 517; DB 2; Length 622;
Best Local Similarity 99.8%; Pred. No. 1.3e-279;
Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY GCAACATGAACCTGGCTGATGCCAGGCAAGGTTTGACATGATTTGACA 546
18 GCAACATGAACCTGGCTGATGCCAGGCAAGGTTTGACATGATTTGACA 77
QY AGCAGTTGAGGCTTATCATGGAATCCATGCTTCCCGGATCTATGACCCGAT 606
78 AGCAGTTGAGGCTTATCATGGAATCCATGCTTCCCGGATCTATGACCCGAT 137
QY TCTTTCCTGCCAACAGAGAGAGCTGCCAACACTCATCAAGGCAACAGAGAGC 666
138 TCTTTCCTGCCAACAGAGAGAGCTGCCAACACTCATCAAGGCAACAGAGAGC 197
QY AAGTTCAACAAATCATCAAGAGATCATGAGGCTTTAAGAGGACCAAGTGAACAAG 726
198 AAGTTCAACAAATCATCAAGAGATCATGAGGCTTTAAGAGGACCAAGTGAACAAG 257
QY TGGTTGACTGTGAGTGCACACAGAGAGTACAGTAATTTGTTGGGCTTAAATG 786
258 TGGTTGACTGTGAGTGCACACAGAGAGTACAGTAATTTGTTGGGCTTAAATG 317
QY AACACATGAGAAATCTTGGCTGTGAGACGAATGAGGCTGAGATTTCTCTTCCA 846
318 AACACATGAGAAATCTTGGCTGTGAGACGAATGAGGCTGAGATTTCTCTTCCA 377
QY CCTTGTATGCAATGCTTGTGATGAAATGTTCTTTTATTAAATGAAAGCCCTCAGA 906
378 CCTTGTATGCAATGCTTGTGATGAAATGTTCTTTTATTAAATGAAAGCCCTCAGA 437
QY ACACTTTTGTACAGGAGCTGATTTGATCTTGCCATGCGAGAACACTTTGATGAG 966
438 ACACTTTTGTACAGGAGCTGATTTGATCTTGCCATGCGAGAACACTTTGATGAG 497
QY ATGACTTTCAAGATGCTGACCAAAATGAAATCTGTTGGTTGATTTCTTGTGGGG 1026
498 ATGACTTTCAAGATGCTGACCAAAATGAAATCTGTTGGTTGATTTCTTGTGGGG 557
QY CTGGTATCAAGCCCAATCTATAGTCAG 1054
558 CTGGTATCAAGCCCAATCTATAGTCAG 585

RESULT 6
BE331363/c
LOCUS
DEFINITION
Gm-c1041-1394 bp mRNA linear EST 13-JUL-2004
8098101.y1 Gm-c1041 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
1-PHOSPHATE SYNTHASE ; mRNA sequence.
ACCESSION
BE331363
VERSION
BE331363.1 GI:9205139
KEYWORDS
EST.

SOURCE
ORGANISM
Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
1 (bases 1 to 619)

AUTHORS

Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
McGunn, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McGunn, R., Waterston, R. and Wilson, R.

TITLE

JOURNAL

COMMENT

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: estewatson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. Possibly
reversed clone: similarly on wrong strand This clone is available
through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA
(phone: 800 423 4163; email: info@biogeneticservices.com)
High quality sequence stop: 491.

FEATURES

source

1..619

/organism="Glycine max"

/mol_type="mRNA"

/cultivar="Williams"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1041-1394"

/tissue_type="Senescing leaves, mature plant, greenhouse
grown"

/lab_host="DH10B"

/clone_lib="Gm-c1041"

/note="Vector: pT73Pac (Pharmacia); Site 1: EcoRI;
Site 2: HindIII; This library was constructed from mRNA
isolated from senescing leaf tissue of mature greenhouse
grown plants of the cultivar Williams. Complementary DNA
was synthesized from mRNA using a 3' anchored poly(dT)
primer. EcoRI adapters were ligated to the blunt-ended
cDNA fragments followed by digestion with EcoRI and
HindIII. The cDNA fragments were directionally cloned
into the EcoRI-HindIII restriction site of the pT73-Pac
vector. The ligated cDNA fragments were transformed into
DH10B host cells. This library was constructed by Dr.
Randy Shoemaker."

ORIGIN

Query Match 28.9%; Score 508; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 1.6e-274;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1252 ATGATGACACTTCAGAGATATTCATGGGTGAAGAGCACTGTTTGGACAACAT 1311
519 ATGATGACACTTCAGAGATATTCATGGGTGAAGAGCACTGTTTGGACAACAT 460
QY 1312 GCGAGGATCCCTCTTAGTGTCTCTTATTTGAGACTTGTCTTGTGAGCTCA 1371
459 GCGAGGATCCCTCTTAGTGTCTCTTATTTGAGACTTGTCTTGTGAGCTCA 400
QY 1372 GCACTGAGATGAGTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1431
399 GCACTGAGATGAGTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 340
QY 1432 CTACCATCTCAGCTACCTCAACAGAGCTCTCTGTTTCAACCGGGTACACAGTGTGA 1491
339 CTACCATCTCAGCTACCTCAACAGAGCTCTCTGTTTCAACCGGGTACACAGTGTGA 280

```
QY 1492 ATGCAATTGTCACAGCAGCGTCGCAATGCTGGAAAACATAATGAGCGCTGTGTGGATTGG 1551
DB 279 ATGCAATTGTCACAGCAGCGTCGCAATGCTGGAAAACATAATGAGCGCTGTGTGGATTGG 220
QY 1552 CCCCAGAAATTAACATGATTTCTCGAGTACAACTGGAAGCATGGACCGAAGATTAATATG 1611
DB 219 CCCCAGAAATTAACATGATTTCTCGAGTACAACTGGAAGCATGGACCGAAGATTAATATG 160
QY 1612 TTGGGGTACCTAGCTGGAATGTTTTATGTTAATATATGTTTGTCTTAATATTTTGCAGT 1671
DB 159 TTGGGGTACCTAGCTGGAATGTTTTATGTTAATATATGTTTGTCTTAATATTTTGCAGT 100
QY 1672 GTAATTGAATGATCATGCTTCATTAATGCTTTAGACGGCGCAATATCTGTTACTAGCA 1731
DB 99 GTAATTGAATGATCATGCTTCATTAATGCTTTAGACGGCGCAATATCTGTTACTAGCA 40
QY 1732 ACATGAATGAATGATGATTAATTTGTG 1759
DB 39 ACATGAATGAATGATGATTAATTTGTG 12
```

```
RESULT 7
AM397453 575 bp mRNA linear EST 14-JUL-2004
LOCUS AM397453
DEFINITION Gm-c1007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
MYO-1NOSITOL-1-PHOSPHATE SYNTHASE ;, mRNA sequence.
ACCESSION AM397453
VERSION AM397453.1 GI:6915923
KEYWORDS EST.
```

```
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 575)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCaum, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1909 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 414.
Location/Qualifiers
```

```
FEATURES
source
1..575
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1007-2594"
/lab_host="DH10B"
/clone_lib="Gm-c1007"
/notes="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants (individual
seed fresh weight of 100-300mg). The library was prepared
using the Life Technologies pSuperScript cDNA library
```

construction kit. Complementary DNA was synthesized from mRNA using a poly (dt) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

Query Match 28.7%; Score 505; DB 2; Length 575;
Beet Local Similarity 100.0%; Pred. No. 8e-273;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 13 CTTTGTAAATTTCTATCATCTTATCTTTGTGAAAAAATATGTTCTCAGAAATTTTA 72
DB 1 CTTTGTAAATTTCTATCATCTTATCTTTGTGAAAAAATATGTTCTCAGAAATTTTA 60
QY 73 AGTTGAATGTCCTAATGATGATGACCGAGACTGATTCAGTCCGTGACACTAGC 132
DB 61 AGTTGAATGTCCTAATGATGATGACCGAGACTGATTCAGTCCGTGACACTAGC 120
QY 133 AAACCCAGAACTGTTTACAGAGAAACAGAAATGCACTATCAGTATGTCAAACCA 192
DB 121 AAACCCAGAACTGTTTACAGAGAAACAGAAATGCACTATCAGTATGTCAAACCA 180
QY 193 AATCTGCAAAATACGAATTTAAACCAACATCATCTTCTAATTTAGGGTAAATGCTTG 252
DB 181 AATCTGCAAAATACGAATTTAAACCAACATCATCTTCTAATTTAGGGTAAATGCTTG 240
QY 253 TGGGTTGGGGTGGAAAACACGCTTACCTCACCCTGAGTGTATTTCTAACCAGAGG 312
DB 241 TGGGTTGGGGTGGAAAACACGCTTACCTCACCCTGAGTGTATTTCTAACCAGAGG 300
QY 313 GCATTTCAATGGGCTACAAAGAACAGATTTCAACAGCAATTTAGTGGTCCCTCACC 372
DB 301 GCATTTCAATGGGCTACAAAGAACAGATTTCAACAGCAATTTAGTGGTCCCTCACC 360
QY 373 AAGCTCAGCTATCCGAGTGGGCTCTTCCAGGGAGAGAAATCATGCCCATTCACA 432
DB 361 AAGCTCAGCTATCCGAGTGGGCTCTTCCAGGGAGAGAAATCATGCCCATTCACA 420
QY 433 GCTTCTTCCAAATGATTAACCTGACGACATTTGTTGGGGGATGGATATACGACA 492
DB 421 GCTTCTTCCAAATGATTAACCTGACGACATTTGTTGGGGGATGGATATACGACA 480
QY 493 TGAACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 517
DB 481 TGAACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 505
```

```
RESULT 8
AT194750 648 bp mRNA linear EST 12-JUL-2004
LOCUS AT194750
DEFINITION sb66f07.y1 Gm-c1019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
MYO-1NOSITOL-1-PHOSPHATE SYNTHASE ;, mRNA sequence.
ACCESSION AT194750
VERSION AT194750.1 GI:5342466
KEYWORDS EST.
```

```
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
```

```
REFERENCE 1 (bases 1 to 648)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
```


TITLE
JOURNAL
COMMENT

McCann, R., Waterson, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. Possible
reversed clones: similarity on wrong strand this clone is available
through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA
(phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1867 Std Error: 0.00
High quality sequence stop: 438.
Location/Qualifiers

FEATURES

source

1..648
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-158"
/issue_type="Immature seed coats of greenhouse grown
plants"
/lab_host="DH10B (Gibco BRL)"
/clone_lib="Gm-cl019"
/note="Vector: pSPORT1 (Life Technologies); Site 1: Not I;
Site 2: Sal I; This cDNA library was constructed from mRNA
isolated from immature seed coats (200-300 mg) of
greenhouse grown plants. The library was prepared using
the Life Technologies pSuperScript cDNA library
construction kit. Complementary DNA was synthesized from
mRNA using a poly (dT) sequence with a Not I restriction
site. Sal I linker adapters were ligated to the
blunt-ended cDNA fragments followed by Not I digestion.
The cDNA fragments were directionally cloned into the Not
I-Sal I restriction site of the pSPORT1 vector. The
ligated cDNA fragments were transformed into E.coli
Electromax DH10B host cells (Gibco BRL). This library was
constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

ORIGIN

Query Match 28.5%; Score 501; DB 1; Length 648;
Best Local Similarity 99.8%; Pred. No. 1.5e-270;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 38 ATCTTGTAAAAAATATGTTTCATCGAATTTTAACTGAGTCTCTTAATGTGAAGTA 97
DB 18 ATCTTGTAAAAAATATGTTTCATCGAATTTTAACTGAGTCTCTTAATGTGAAGTA 77
QY 98 CACCGAGACTGAGATTCAGTCCGTGTACAACTACGAAACACCGAAGCTTGTACAGAGAA 157
DB 78 CACCGAGACTGAGATTCAGTCCGTGTACAACTACGAAACACCGAAGCTTGTACAGAGAA 137
QY 158 CAGGAATGACCTTACAGTGTGTCAAAACCAATCTGTCAATATACGAATTTTAAAC 217
DB 138 CAGGAATGACCTTACAGTGTGTCAAAACCAATCTGTCAATATACGAATTTTAAAC 197
QY 218 CAACATCCATGTTCTTAATTAAGGGTATGCTTGTGGGTGGGAGAAACAAGGCTC 277
DB 198 CAACATCCATGTTCTTAATTAAGGGTATGCTTGTGGGTGGGAGAAACAAGGCTC 257
QY 278 AACCTTACACCGGTGTGTATTGCTACACGAGAGGAGCATTTTCATGGGCTACAAAGACAA 337
DB 258 AACCTTACACCGGTGTGTATTGCTACACGAGAGGAGCATTTTCATGGGCTACAAAGACAA 317
QY 338 GATTCAACAAGCCAAATTAATTGGCTCCCTACCCAGGCTCAGCTATCCGAGTTGGGTC 397
DB 318 GATTCAACAAGCCAAATTAATTGGCTCCCTACCCAGGCTCAGCTATCCGAGTTGGGTC 377
QY 398 CTTCCAGGAGAGGAATCTATGCCCATTTCAAGAGCCCTGCTTCAATGTGTTAACTCTGA 457

DB 378 CTTCCAGGAGAGGAATCTATGCCCATTTCAAGAGCCCTGCTTCAATGTGTTAACTCTGA 437
QY 458 CCAATTTGTTTGGGGAGATGAGATATACGAACATGAACCTGGCTGATGCCATGGCCAG 517
DB 438 CCAATTTGTTTGGGGAGATGAGATATACGAACATGAACCTGGCTGATGCCATGGCCAG 497
QY 518 GCGAAAGCTTTGATCATGATTTGACAGAGAGAGTGGAGCCCTTACATGAATCCATGGCT 577
DB 498 NCCAAAGGTTTGCATCATGATTTGCAAGAGAGTGGAGCCCTTACATGAATCCATGGCT 557
QY 578 TCCACTCCCCCG 589
DB 558 TCCACTCCCCCG 569

RESULT 9
AM317927
LOCUS
DEFINITION
G558h08.y1 Gm-cl007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl007-664 5' similar to FR:040271 Q40271
MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AM317927
GI:6747471
EST.

Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (baaes 1 to 539)
Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Treising, B., Allen, M.,
Bowers, Y., Persson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterson, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 432.
Location/Qualifiers

FEATURES

source

1..539
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl007-664"
/lab_host="DH10B"
/clone_lib="Gm-cl007"
/note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants (individual
seed fresh weight of 100-300mg). The library was prepared
using the Life Technologies pSuperScript cDNA library
construction kit. Complementary DNA was synthesized from
mRNA using a poly (dT) sequence with a Not I restriction
site. Sal I linker adapters were ligated to the
blunt-ended cDNA fragments followed by Not I digestion. The
cDNA fragments were directionally cloned into the
NotI-SalI restriction site of the pSPORT1 vector. The

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 478)
Periapuram, C.C., Li, L., Wurtele, E.S., Westgate, M.E. and Nikolau, B.J.
Molecular determination of soybean composition
Unpublished (2004)
Contact: Basil J. Nikolau and Cyril C. Periapuram
Department of Biochemistry, Biophysics and Molecular Biology
Iowa State University
210 Molecular Biology Building, Ames, IA 50011, USA
Tel: 515 294 9423
Fax: 515 294 0453
Email: dimmas@iastate.edu

Individual base call and confidence value were assigned using the Phred software (http://www.phrap.org/). Overall sequence quality assessment and vector trimming were conducted using the Lucy software (http://www.tigr.org/software/).
This clone was originally generated by the Public Soybean EST Project (http://129.186.26.94/soybeanest.html/Shoemaker, R (rcshoe@iastate.edu)).
This clone is available through: Biogenetic Services, 801 32nd Ave, Brookings, SD 57006. For further information call 605-697-8500 or contact info@biogeneticservices.com
Seq primer: T7-1 (5' AAT ACG ACT CAC TAT AG 3').

FEATURES
source Location/Qualifiers

1..478
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1030-2818"
/lab_host="DH10B"
/clone_1lb="Gm-r1030"
/note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; This cDNA library was constructed from mRNA isolated from immature cotyledons of greenhouse grown plants (individual seed fresh weight of 100-300mg). The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restriction site. SalI linker adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed by Dr. Lilia Vokkin and Dr. Ann Khanna. Note that Gm-r1030 is a re-rack of Gm-cl1007."

ORIGIN

Query Match 26.5%; Score 467; DB 7; Length 478;
Best Local Similarity 100.0%; Pred. No. 2.3e-251;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OR 19 GGAATTCATTCATTTTATCTTTGTGAAAAATATGTTTCATCGAAATTTTAAAGCTTG 78
DB 12 GGAATTCATTCATTTTATCTTTGTGAAAAATATGTTTCATCGAAATTTTAAAGCTTG 71
OR 79 AGTGTCTTATGTAAGTACACGAGATCGAGTTCAGTCCGGTACCACTCGAAGACA 138
DB 72 AGTGTCTTATGTAAGTACACGAGATCGAGTTCAGTTCGCGTACCACTCGAAGACA 131
OR 139 CGGAATCTGTCAGGAGACAGGAATGCACTATCAGTGAATGTCGAAACCAATCTTG 198
DB 133 CGGAATCTGTCAGGAGACAGGAATGCACTATCAGTGAATGTCGAAACCAATCTTG 191
OR 199 TCAATACGAATTTAAACCAACATCAGTTCCTAAATTTAGGGGTAATGCTTGGGTT 258
DB 192 TCAATACGAATTTAAACCAACATCAGTTCCTAAATTTAGGGGTAATGCTTGGGTT 251
OR 259 GGGGTGAAACAACGAGCTCAACCGGTGGTGTATGCTAACGAGAGGGCATTT 318

DB 252 GGGGTGAAACAACGAGCTCAACCGGTGGTGTATGCTAACGAGAGGGCATTT 311
OR 319 CATGGCTACAAAGACAAAGATTCAACAGCCATTTAGCTCCCTCAACCAAGCCT 378
DB 312 CATGGCTACAAAGACAAAGATTCAACAGCCATTTAGCTCCCTCAACCAAGCCT 371
OR 379 CAGCTATCCGAGTGGGTGCTTCACAGAGAGGAATTCATGCCCATTCAAGAGCTGC 438
DB 372 CAGCTATCCGAGTGGGTGCTTCACAGAGAGGAATTCATGCCCATTCAAGAGCTGC 431
OR 439 TTCCAAATGTTAAACCTGACGACATTTGTTGGGGAATGCGATATC 485
DB 432 TTCCAAATGTTAAACCTGACGACATTTGTTGGGGAATGCGATATC 478

RESULT 12
BO612082 443 bp mRNA linear EST 05-JUL-2004
LOCUS sap7b07.y1 Gm-cl087 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl087-6590.5' similar to TR:Q9SSV4 Q9SSV4 MYO-INOSITOL-1-PHOSPHATE SYNTHASE.; mRNA sequence.
DEFINITION
ACCESSION BO612082 GI:21601751
VERSION BO612082
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE

1 (bases 1 to 443)
Shoemaker, R., Keim, P., Vokkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schunk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

JOURNAL

Public Soybean EST Project

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 415.
Location/Qualifiers

FEATURES
source

1..443
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl087-6590"
/issue_type="Roots"
/lab_host="DH10B"
/clone_1lb="Gm-cl087"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The mRNA was prepared using polyactract mRNA system kit from PROMEGA. The cDNA was prepared using the STRATAGENE kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site
(5'GAGAGAGAGAGAGAGAGAACTGCTGACTTTTATTTTATTTT).
EcoRI adapters (5'OH-AATTCGACAGAG and 3'GCCGCTCCTCp) were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction sites of the pBluescript

vector. The vector was previously dephosphorylated. The ligated cDNA fragments were transformed into DH10B by electroporation. Library is in LB medium with ampicillin and glycerol 8%. Average insert size: 800 bp."

ORIGIN

Query Match 25.2%; Score 443; DB 5; Length 443;

Best Local Similarity 100.0%; Pred. No. 8.2e-238; Mismatches 0; Indels 0; Gaps 0;

Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

854 TGCATTCCTTGTATGAAATGTTCTTTCAATTAAGAGCCCTCAGAACACTTT 913
1 TCCCATTCCTTGTATGAAATGTTCTTTCAATTAAGAGCCCTCAGAACACTTT 60
914 TTTACCAAGGCTGATTTGATCTTCCATCGCGAGAACACTTTGATGAGATGACTT 973
61 TGTACCAAGGCTGATTTGATCTTCCATCGCGAGAACACTTTGATGAGATGACTT 120

974 CAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTTTGGGGCTGGTAT 1033
121 CAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTTTGGGGCTGGTAT 180

1034 CAAGCCACATCTAATAGTACATTCACATCTGGGAAACAATGATGATGATCTTTC 1093
181 CAAGCCACATCTAATAGTACATTCACATCTGGGAAACAATGATGATGATCTTTC 240

1094 GGCTCCACAACTTTCCGTTCCAGGAATCTTCCAGACACGTTTGTGATGATGAT 1153
241 GGCTCCACAACTTTCCGTTCCAGGAATCTTCCAGACACGTTTGTGATGATGAT 300

1154 CAACAGCATGCGATCTCTATGAGCTGGTGAACATCCAGACCAATGTTGTTATTTAA 1213
301 CAACAGCATGCGATCTCTATGAGCTGGTGAACATCCAGACCAATGTTGTTATTTAA 360

1214 GATATGCTTTCATGTAAGGAGACAGCAAGACCATGATGATGATGATGATGAT 1273
361 GATATGCTTTCATGTAAGGAGACAGCAAGACCATGATGATGATGATGATGAT 420

1274 CATGGTGGAAAGACGACCATTTG 1296
421 CATGGTGGAAAGACGACCATTTG 443

RESULT 13
BI347339 420 bp mRNA linear EST 30-OCT-2001
LOCUS F165B19 cDNA library from Forrest roots infected by F. solani
DEFINITION Glycine max cDNA clone F165B19, mRNA sequence.
ACCESSION BI347339
VERSION BI347339.1 GI:16519554
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 420)
Iqbal,M.J. and Lightfoot,D.A.
Differentially expressed mRNA sequences identified by differential display of mRNA from soybean, Glycine max (L.) Merr. var. Forrest, roots in response to Fusarium solani f. sp. glycines inoculation
Unpublished (2001)
Contact: Iqbal MJ and Lightfoot, DA
Center of Excellence in Soybean Research, Teaching and Outreach
Southern Illinois University at Carbondale
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA
Tel: 618 453-3121
Fax: 618 453-7457
Email: mj1qbal@siu.edu, ga4082@siu.edu
Seq primer: M13 Forward.
Location/Qualifiers

FEATURES

source

1..420
/organism="Glycine max"
/mol_type="mRNA"
/cultiivar="Forrest"
/db_xref="taxon:3847"
/clone="F165B19"
/dev_stage="14 days after inoculation with F. solani"
/clone_11bs="cDNA library from Forrest roots infected by F. solani"
/note="Organ: Root; Vector: pGEM; Forrest roots were inoculated with Fusarium solani f. sp. glycines and samples were collected after 14 days of inoculation. Total RNA was used for cDNA synthesis using SMART PCR cDNA synthesis kit (CLONTECH, CA) and cloned in pGEM vector. ESTs were sequenced using M13 universal forward or reverse primer."

ORIGIN

Query Match 22.7%; Score 400; DB 4; Length 420;

Best Local Similarity 100.0%; Pred. No. 1.6e-213; Mismatches 0; Indels 0; Gaps 0;

Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTCTTCTTATTCCTTTGTAATTCATTCATCTTAACTTTGGAATAATGTTCA 60
21 CTCTTCTTATTCCTTTGTAATTCATTCATCTTAACTTTGGAATAATGTTCA 80
61 TCGAATTTTAAAGGTGAGTCTCTAATGTAAGTACACGAGACTGAGATTCAGTCCG 120
81 TCGAATTTTAAAGGTGAGTCTCTAATGTAAGTACACGAGACTGAGATTCAGTCCG 140

121 TGTACACTACGAAACCAACGAACTTGTTCAGAGAACGAAATGGACCTATCAGTGA 180
141 TGTACACTACGAAACCAACGAACTTGTTCAGAGAACGAAATGGACCTATCAGTGA 200

181 TTGTAAACCCAAATCTGTCAATTCGAATTTAAACCAATCCATGTCCTAATAG 240
201 TTGTAAACCCAAATCTGTCAATTCGAATTTAAACCAATCCATGTCCTAATAG 260

241 GGCTATGCTTGTGGGTTGGGGTGGAAACAACGCTCAACCTCAGCGTGTGTTATG 300
261 GGCTATGCTTGTGGGTTGGGGTGGAAACAACGCTCAGCGTGTGTTATG 320

301 CTACCCGAGAGGACATTTATGAGGCTACAAAGACAAATTCACAAAGCAATTAATTG 360
321 CTACCCGAGAGGACATTTATGAGGCTACAAAGACAAATTCACAAAGCAATTAATTG 380

361 GCTCCCTCACCAAGCCTCAGCTATCCAGTTGGGTCCTT 400
381 GCTCCCTCACCAAGCCTCAGCTATCCAGTTGGGTCCTT 420

RESULT 14
BE609839 487 bp mRNA linear EST 13-UTL-2004
LOCUS sq45a10.y1 Gm-c1019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1019-8371 5' similar to TR:064437 064437 MVO-INOSITOL PHOSPHATE SYNTHASE. ;, mRNA sequence.
ACCESSION BE609839
VERSION BE609839.1 GI:9900871
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 487)
Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. Trace considered overall poor quality. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 1.

FEATURES

source

1. 487
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1019-8371"
/issue_type="Immature seed coats of greenhouse grown plants"
/lab_host="DH10B (Gibco BRL)"
/clone_lib="Gm-c1019"
/note="Vector: pSPORT1 (Life Technologies); Site 1: Not I; Site 2: Sal I; This cDNA library was constructed from mRNA isolated from immature seed coats (200-300 mg) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells (Gibco BRL). This library was constructed by Dr. Lila Vodka and Dr. Ann Khanna."

ORIGIN

Query Match 22.2%; Score 391; DB 2; Length 487;

Best Local Similarity 99.8%; Pred. No. 2e-208; 1; Indels 0; Gaps 0;

Matches 441; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

623 AGAGAGCGTGCACAAAGCTCATCAAGGCGCAAAAGAGCAAGTCAACAATCAT 682
1 AGAGAGCGTGCACAAAGCTCATCAAGGCGCAAAAGAGCAAGTCAACAATCAT 60
683 CAAAGCATCAAGCGCTTTAAGAGACCAAAAGTGAACAAGTGTGTACTGTGAC 742
61 CAAAGCATCAAGCGCTTTAAGAGACCAAAAGTGAACAAGTGTGTACTGTGAC 120
743 TCCCAACAGAGAGGTACATTAATTTGGTGGCCCTTAATGACACATGAGAAATCT 802
121 TCCCAACAGAGAGGTACATTAATTTGGTGGCCCTTAATGACACATGAGAAATCT 180
803 CTGGGCTGTGAGACAGAAATGAGGCTGAGATTTCTCCTTCACCTTGATGCCATTGC 862
181 CTGGGCTGTGAGACAGAAATGAGGCTGAGATTTCTCCTTCACCTTGATGCCATTGC 240
863 TTGTGTTATGAAATGTTCTTTTCATTAATGAGAGCCCTCAGAACACTTTGTACAG 922
241 TTGTGTTATGAAATGTTCTTTTCATTAATGAGAGCCCTCAGAACACTTTGTACAG 300
923 GCTGATGATCTTGGCATGCGAGAGAACCTTTGATTTGGTGGAGATGACTTCAAGGTGG 982
301 GCTGATGATCTTGGCATGCGAGAGAACCTTTGATTTGGTGGAGATGACTTCAAGGTGG 360
983 TCAGACCAAAATGAATCTGTGTGTTGATTTCTTTGGGGGCTGTATCAAGCCAAC 1042
|||||

Db 361 TCAGACCAAAATGAATCTGTGTGTTGATTTCTTTGGGGGCTGTATCAAGCCAAC 420
Qy 1043 ATCTATGTCAGTTACCAACAT 1064
Db 421 ATCTATGTCAGTTACCAACAT 442

RESULT 15

LOCUS AM472457 435 bp mRNA linear EST 14-JUL-2004
DEFINITION s125e04.y1 Gm-c1029 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1029-2263 5' similar to SW:IN01 SPIPO P42803

MYO-INOSITOL-1-PHOSPHATE SYNTHASE ; mRNA sequence.
ACCESSION AM472457
VERSION AM472457.1 GI:7042563

KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 435)

REFERENCE Shoemaker,R., Keim,P., Vodka,L., Erpelting,J., Corryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

TITLE

JOURNAL

COMMENT

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1108 Std Error: 0.00.

Location/Qualifiers

FEATURES

source

1. 435
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1029-2263"
/issue_type="very young cotyledons of greenhouse grown plants"
/lab_host="DH10B"
/clone_lib="Gm-c1029"
/note="Vector: pSPORT1; Site 1: Not I; Site 2: Sal I; This cDNA library was constructed from mRNA isolated from very young cotyledons (20-50mg fresh weight) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed by Dr. Paul Keim and Dr. Virginia Corryell."

ORIGIN

Query Match 21.8%; Score 383; DB 2; Length 435;

Best Local Similarity 99.8%; Pred. No. 6.4e-204; 1; Indels 0; Gaps 0;

Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY 698 GTTTAAGAAGCCACCAAGTGAACAAGTGGTTGTACTGTGGACTGCGCAACAGAGAG 757
    |||||||
Db 2 GTTTAAGGAAGCCACCAAGTGAACAAGTGGTTGTACTGTGGACTGCGCAACAGAGAG 61
    |||||||
OY 758 GTACAGTAATTTGGTTGGGCTTATGACACACATGAGAAATCTTGGCTGTGTGA 817
    |||||||
Db 62 GTACAGTAATTTGGTTGGGCTTATGACACACATGAGAAATCTTGGCTGTGTGA 121
    |||||||
OY 818 CAGAAATGAGGCTGAGATTCTCCTCCACCTTGATGCAATTGCTTGTGTATGGA 877
    |||||||
Db 122 CAGAAATGAGGCTGAGATTCTCCTCCACCTTGATGCAATTGCTTGTGTATGGA 181
    |||||||
OY 878 TGTTCCTTCATTTAATGAAGCCCTCAGAACATTTTGTACAGGGCTGATGATCTTGC 937
    |||||||
Db 182 TGTTCCTTCATTTAATGAAGCCCTCAGAACATTTTGTACAGGGCTGATGATCTTGC 241
    |||||||
OY 938 CATCGGAGGAACTTTGATTGGTGAATGACCTTCAAGAGTGTGAGACCAAAATGAA 997
    |||||||
Db 242 CATCGGAGGAACTTTGATTGGTGAATGACCTTCAAGAGTGTGAGACCAAAATGAA 301
    |||||||
OY 998 ATCTGTGTTGGTTGATTTCTGTGGGGCTGGTATCAAGCCAATCTATAGTCAGTTA 1057
    |||||||
Db 302 ATCTGTGTTGGTTGATTTCTGTGGGGCTGGTATCAAGCCAATCTATAGTCAGTTA 361
    |||||||
OY 1058 CAACCATCTGGGAAACAATGATGATGATCTTTGGCTCCACAACCTTCCGTTCCAA 1117
    |||||||
Db 362 CAACCATCTGGGAAACAATGATGATGATCTTTGGCTCCACAACCTTCCGTTCCAA 421
    |||||||
OY 1118 GGAATCTCCAGA 1131
    |||||||
Db 422 GGAATCTCCAGA 435
    |||||||
```

Search completed: June 8, 2005, 14:33:44
Job time : 6002.12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: June 8, 2005, 06:50:16 ; Search time 299.302 Seconds
(without alignments)
9621.891 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttctctctcttctgt.....aatgtagtataatttctgt 1760

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
 - 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
 - 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
 - 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
 - 5: /cgn2_6/prodata/1/ina/5C_COMB.seq:*
 - 6: /cgn2_6/prodata/1/ina/5D_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	1.3	1931	3	US-09-118-442-10
2	23	1.3	1931	3	US-09-677-064-10
3	23	1.3	3546	3	US-09-118-442-14
4	23	1.3	3546	3	US-09-118-442-15
5	23	1.3	3546	3	US-09-677-064-14
6	23	1.3	3546	3	US-09-677-064-15
7	20	1.1	27223	4	US-09-949-016-13036
8	20	1.1	139049	4	US-09-949-016-17030
9	19	1.1	152132	4	US-09-949-016-13845
10	19	1.1	152132	4	US-09-949-016-12371
11	19	1.1	152132	4	US-09-949-016-12371
12	19	1.1	152132	4	US-09-949-016-12371
13	19	1.1	152132	4	US-09-949-016-12371
14	19	1.1	152132	4	US-09-949-016-12371
15	19	1.1	152132	4	US-09-949-016-12371
16	19	1.1	152132	4	US-09-949-016-12371
17	19	1.1	152132	4	US-09-949-016-12371
18	19	1.1	152132	4	US-09-949-016-12371
19	19	1.1	152132	4	US-09-949-016-12371
20	19	1.1	152132	4	US-09-949-016-12371
21	19	1.1	152132	4	US-09-949-016-12371
22	19	1.1	152132	4	US-09-949-016-12371
23	19	1.1	152132	4	US-09-949-016-12371
24	19	1.1	152132	4	US-09-949-016-12371
25	19	1.1	152132	4	US-09-949-016-12371
26	19	1.1	152132	4	US-09-949-016-12371
27	19	1.1	152132	4	US-09-949-016-12371

28	18	1.0	1326	4	US-09-134-000C-3096	Sequence 3096, App
29	18	1.0	1329	3	US-09-071-035-313	Sequence 313, App
30	18	1.0	2219	3	US-08-510-646B-17	Sequence 17, App1
31	18	1.0	2477	4	US-09-907-794A-169	Sequence 169, App
32	18	1.0	2477	4	US-09-905-125A-169	Sequence 169, App
33	18	1.0	2477	4	US-09-902-775A-169	Sequence 169, App
34	18	1.0	2477	4	US-09-906-700-169	Sequence 169, App
35	18	1.0	2477	4	US-09-903-603A-169	Sequence 169, App
36	18	1.0	2477	4	US-09-904-920A-169	Sequence 169, App
37	18	1.0	2477	4	US-09-909-064-169	Sequence 169, App
38	18	1.0	2477	4	US-09-905-381A-169	Sequence 169, App
39	18	1.0	2477	4	US-09-906-618-169	Sequence 169, App
40	18	1.0	4989	4	US-09-693-011-112	Sequence 12, App1
41	18	1.0	5083	4	US-09-693-011-11	Sequence 11, App1
42	18	1.0	6314	4	US-09-693-011-9	Sequence 10, App1
43	18	1.0	6408	4	US-09-693-011-9	Sequence 9, App1
44	18	1.0	16573	4	US-09-949-016-11764	Sequence 11764, A
45	18	1.0	17370	4	US-09-949-016-17331	Sequence 17331, A

ALIGNMENTS

```

RESULT 1
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10

Query Match 1.3%; Score 23; DB 3; Length 1931;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1158 AGCAATGCCATCTCTATGAGCC 1180
Db 1203 AGCAATGCCATCTCTATGAGCC 1225

RESULT 2
US-09-677-064-10
; Sequence 10, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064

```



```
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ PRIOR FILING DATE: 1998-07-17
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 10
/ LENGTH: 1931
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-677-064-10

Query Match      1.3%; Score 23; DB 3; Length 1931;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1158 AGCAATGCCATCCTCTATGAGCC 1180
DB      1203 AGCAATGCCATCCTCTATGAGCC 1225

RESULT 3
US-09-118-442-14
/ Sequence 14, Application US/09118442B
/ Patent No. 6197561
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ APPLICANT: Wang, Xun
/ APPLICANT: Bowen, Benjamin A.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ FILE REFERENCE: 0706
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ CURRENT FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 14
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-118-442-14

Query Match      1.3%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1158 AGCAATGCCATCCTCTATGAGCC 1180
DB      2842 AGCAATGCCATCCTCTATGAGCC 2864

RESULT 4
US-09-118-442-15
/ Sequence 15, Application US/09118442B
/ Patent No. 6197561
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ APPLICANT: Wang, Xun
```

```
/ APPLICANT: Bowen, Benjamin A.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ FILE REFERENCE: 0706
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ CURRENT FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-118-442-15

Query Match      1.3%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1158 AGCAATGCCATCCTCTATGAGCC 1180
DB      2842 AGCAATGCCATCCTCTATGAGCC 2864

RESULT 5
US-09-677-064-14
/ Sequence 14, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677,064
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ PRIOR FILING DATE: 1998-07-17
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 14
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-677-064-14

Query Match      1.3%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1158 AGCAATGCCATCCTCTATGAGCC 1180
DB      2842 AGCAATGCCATCCTCTATGAGCC 2864

RESULT 6
US-09-677-064-15
/ Sequence 15, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
```

```

; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; FILE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-15

```

```

Query Match 1.3%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1158 AGCAATGCCATCCTCTATGAGCC 1180
DB 2842 AGCAATGCCATCCTCTATGAGCC 2864

```

```

RESULT 7
US-09-949-016-13036
; Sequence 13036, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13036
; LENGTH: 27223
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13036

```

```

Query Match 1.1%; Score 20; DB 4; Length 27223;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1529 AATGAGGCTGTGTGTGAT 1548
DB 3336 AATGAGGCTGTGTGTGAT 3355

```

```

RESULT 8
US-09-949-016-17030/c
; Sequence 17030, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17030
; LENGTH: 139049
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(139049)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17030

```

```

Query Match 1.1%; Score 20; DB 4; Length 139049;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 229 TTCCTAATTAGGGTATG 248
DB 30577 TTCCTAATTAGGGTATG 30558

```

```

RESULT 9
US-09-175-928-9/c
; Sequence 9, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3153
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-9

```

```

Query Match 1.1%; Score 19; DB 3; Length 3153;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1153 TCACAGCATGCCATCCT 1171
DB 178 TCACAGCATGCCATCCT 160

```

```

RESULT 10
US-09-949-016-13845
; Sequence 13845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

```

: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13845
: LENGTH: 152132
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-13845

```

Query Match	1.1%;	Score 19;	DB 4;	Length 152132;
Best Local Similarity	100.0%;	Pred. NO. 48;		
Matches 19;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	860	TGCTTGTATTGAAAT	878
Db	130119	TGCTTGTATTGAAAT	130137

```

RESULT 11
US-09-949-016-12371
; Sequence 12371, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12371
; LENGTH: 152145
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12371

```

	Query Match	1.1%	Score 19	DB 4	Length 152145
	Best Local Similarity	100.0%	Pred. No. 48		
Matches	19	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
Qy	860	TCCTTGCTGTATGCAAAAT	878		
DB	130119	TCCTTGCTGTATGCAAAAT	130137		

```

RESULT 12
US-09-949-016-15868
; Sequence 15868, Application US/09949016
; Patent No. 681239
;
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
;
; CURRENT APPLICATION NUMBER: US/09/949, 016
;
; CURRENT FILING DATE: 2000-04-14
;
; PRIOR APPLICATION NUMBER: 60/241,755
;

```

```

? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0.
? SEQ ID NO 15868
? LENGTH: 374159
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-15868

```

```
Query Match      1.1%; Score 19; DB 4; Length 374159;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	755	GAGGTACAGTAATTGGTT	773
Db	290127	GAGGTACAGTAATTGGTT	290145

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

NAME/KEY: misc_feature
LOCATION: (103998) .. (103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948) .. (148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385) .. (163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989) .. (191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191985) .. (191985)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980) .. (231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187) .. (234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220) .. (234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814) .. (234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398) .. (309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418) .. (309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837) .. (312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312593) .. (312593)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226) .. (319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559157) .. (559157)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241) .. (559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992) .. (600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708) .. (622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081) .. (657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203) .. (657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435) .. (674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442) .. (682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652) .. (713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684) .. (741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature

```

```

LOCATION: (779455) .. (779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676) .. (779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539) .. (855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619) .. (871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830) .. (1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846) .. (1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881) .. (1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881) .. (1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988) .. (1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224) .. (1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473) .. (1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491) .. (1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091) .. (1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020) .. (1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912) .. (1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734) .. (1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998) .. (1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664855) .. (1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 1.1%; Score 19; DB 4; Length 1664976;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 TGTATGGAATGTCT 884
Db 326675 TGTATGGAATGTCT 326693

RESULT 14
US-09-692-570-1
Sequence 1, Application US/09692570
Patent No. 6797466
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
Patent No. 6797466
TITLE OF INVENTION: jannaschii

```

```
/ FILE REFERENCE: PB275C1
/ CURRENT APPLICATION NUMBER: US/09/692,570
/ CURRENT FILING DATE: 2003-01-14
/ PRIOR APPLICATION NUMBER: US 60/024,428
/ PRIOR FILING DATE: 1996-08-22
/ PRIOR APPLICATION NUMBER: US 08/916,421
/ PRIOR FILING DATE: 1997-08-22
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 1
/ LENGTH: 1664976
/ TYPE: DNA
/ ORGANISM: Methanococcus jannaschii
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (28222)..(28222)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (28257)..(28258)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (84773)..(84773)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (84808)..(84808)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (84812)..(84812)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98120)..(98120)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98159)..(98159)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98239)..(98239)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98266)..(98266)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98343)..(98343)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (103998)..(103998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (148948)..(148948)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (163385)..(163385)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (191989)..(191989)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (191995)..(191995)
/ OTHER INFORMATION: n equals a, t, c, or g
```

```
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (231980)..(231980)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234187)..(234187)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234220)..(234220)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309398)..(309398)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309418)..(309418)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312837)..(312837)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312993)..(312993)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (319226)..(319226)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559167)..(559167)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559241)..(559241)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (600992)..(600992)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (622708)..(622708)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (657081)..(657081)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (657203)..(657203)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (674435)..(674435)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (682442)..(682442)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (713652)..(713652)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
```

```

; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

Query Match 1.1%; Score 19; DB 4; Length 1664976;
Best Local Similarity 100.0%; Pred.No. 57;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 866 TGTATGAAATGTTCT 884
DB 326675 TGTATGAAATGTTCT 326693

```

```

RESULT 15
US-09-313-294A-2055
; Sequence 2055, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Ialigudi, Raghnath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 2055
; LENGTH: 256

```

```

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551970H1
US-09-313-294A-2055

```

```

Query Match 1.0%; Score 18; DB 4; Length 256;
Best Local Similarity 100.0%; Pred.No. 96;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 488 CAACATGAACCTGGCTGA 505
DB 99 CAACATGAACCTGGCTGA 116

```

```

Search completed: June 8, 2005, 14:43:08
Job time : 303.302 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 8, 2005, 11:26:41, Search time 1091.38 Seconds
(without alignments)
9956.820 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cccctctctatcccttctgt.....aatgtagtataatttctgt 1760

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 6046767 seqs, 3099530249 residues

Word size: 0

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1760	100.0	1760	14	US-10-025-003-1
2	1760	100.0	1760	14	US-10-718-952-1
3	1538	87.4	1538	18	US-10-424-599-70167
4	1533	87.1	1533	14	US-10-025-003-9
5	1533	87.1	1533	14	US-10-718-952-9
6	1482	84.2	1533	14	US-10-025-003-5
7	1482	84.2	1533	19	US-10-718-952-5

8	295	16.8	594	18	US-10-424-599-70166	Sequence 70166, A
9	276	15.7	661	18	US-10-424-599-116597	Sequence 116597, A
10	146	8.3	634	18	US-10-424-599-22663	Sequence 22663, A
11	146	8.3	1533	14	US-10-025-003-11	Sequence 11, Appl
12	146	8.3	1533	14	US-10-025-003-13	Sequence 13, Appl
13	146	8.3	1533	14	US-10-025-003-15	Sequence 15, Appl
14	146	8.3	1533	14	US-10-718-952-11	Sequence 11, Appl
15	146	8.3	1533	19	US-10-718-952-13	Sequence 13, Appl
16	146	8.3	1533	19	US-10-718-952-15	Sequence 15, Appl
17	146	8.3	2582	18	US-10-424-599-12021	Sequence 12021, A
18	121	6.9	844	18	US-10-424-599-70162	Sequence 70162, A
19	121	6.9	2018	18	US-10-424-599-12022	Sequence 12022, A
20	50	2.8	2582	18	US-10-424-599-12021	Sequence 12021, A
21	48	2.7	989	18	US-10-424-599-12019	Sequence 12019, A
22	47	2.7	355	18	US-10-424-599-38648	Sequence 38648, A
23	35	2.0	136	18	US-10-424-599-13935	Sequence 13935, A
24	30	1.7	553	18	US-10-424-599-78882	Sequence 78882, A
25	30	1.7	553	18	US-10-424-599-83330	Sequence 83330, A
26	29	1.6	377	11	US-09-733-627A-2381	Sequence 2381, Ap
27	29	1.6	566	19	US-10-021-323-12655	Sequence 12655, A
28	29	1.6	618	19	US-10-021-323-17771	Sequence 17771, A
29	28	1.6	365	18	US-10-424-599-93575	Sequence 93575, A
30	28	1.6	262	18	US-10-424-599-94906	Sequence 94906, A
31	26	1.5	252	18	US-10-424-599-12017	Sequence 12017, A
32	26	1.5	1673	20	US-10-425-115-47180	Sequence 47180, A
33	26	1.5	1941	20	US-10-739-930-2091	Sequence 2091, Ap
34	24	1.4	35	14	US-10-025-003-3	Sequence 3, Appl
35	24	1.4	35	14	US-10-718-952-3	Sequence 3, Appl
36	24	1.4	1924	19	US-10-767-701-14714	Sequence 14714, A
37	23	1.3	260	18	US-10-424-599-36573	Sequence 36573, A
38	23	1.3	290	19	US-10-437-963-36182	Sequence 36182, A
39	23	1.3	345	11	US-09-732-627A-3508	Sequence 3508, Ap
40	23	1.3	429	18	US-10-424-599-52735	Sequence 52735, A
41	23	1.3	457	17	US-10-341-961A-259	Sequence 259, App
42	23	1.3	557	19	US-10-021-323-12840	Sequence 12840, A
43	23	1.3	601	19	US-10-021-323-15424	Sequence 15424, A
44	23	1.3	605	19	US-10-021-323-15202	Sequence 15202, A
45	23	1.3	1998	18	US-10-425-114-7081	Sequence 7081, Ap

ALIGNMENTS

RESULT 1
US-10-025-003-1
; Sequence 1, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelt, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-1
Query Match 100.0%; Score 1760; DB 14; Length 1760;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTCTCTTATATCTTTTGTATATTCATTCATCTTAAATCTTTGTGAAAAATATATGTCA 60
1 CTTCTTTTATCTTTTGTATATTCATTCATCTTAAATCTTTGTGAAAAATATATGTCA 60
61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAATACCGAGACTGAGATTCAGTCCG 120
61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAATACCGAGACTGAGATTCAGTCCG 120
121 TGTACAACTAGCAAAACCAACCACTTGTTCACGAGAACGGAATGGACCTATCAGTGA 180
121 TGTACAACTAGCAAAACCAACCACTTGTTCACGAGAACGGAATGGACCTATCAGTGA 180
181 TTGTCAAAACCAAAATCTGTCAAAATGAATTTAAAAACCAATCCATGTTCTTAATAG 240
181 TTGTCAAAACCAAAATCTGTCAAAATGAATTTAAAAACCAATCCATGTTCTTAATAG 240
241 GGGTAACTGTGGGTGGGTGGGAGAACAGGCTCAACCTCAACGGTGTGTTATTTG 300
241 GGGTAACTGTGGGTGGGTGGGAGAACAGGCTCAACCTCAACGGTGTGTTATTTG 300
301 CTTAACCGAGAGGGGCAATTCATGGGCTAACAGAGAACAGATTCAGAACGCAATTA 360
301 CTTAACCGAGAGGGGCAATTCATGGGCTAACAGAGAACAGATTCAGAACGCAATTA 360
361 GCTCCCTACCCCAAGCCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAGAAATCTATG 420
361 GCTCCCTACCCCAAGCCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAGAAATCTATG 420
421 CCCCAATCAAGAGCCTGCTTCCAAATGTGTTAACCTGTGAGACATTTGTTGGGGATGG 480
421 CCCCAATCAAGAGCCTGCTTCCAAATGTGTTAACCTGTGAGACATTTGTTGGGGATGG 480
481 ATATAGCAACATGAACCTGGCTGATGGCATGGCCAGGAGCAAGGTTTGAATCGAT 540
481 ATATAGCAACATGAACCTGGCTGATGGCATGGCCAGGAGCAAGGTTTGAATCGAT 540
541 TGCAGAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCATCCCGGAATCTATGACC 600
541 TGCAGAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCATCCCGGAATCTATGACC 600
601 CGGATTTTCAATGCTGCTCAACCAAGAGAGGCTGCCAACAGCTCATCAAGGGCAACAAAGC 660
601 CGGATTTTCAATGCTGCTCAACCAAGAGAGGCTGCCAACAGCTCATCAAGGGCAACAAAGC 660
661 AAGAGCAAGTTCAAAATATCATCAAGAGACATCAAGGCTTTAAAGAGGCAACAAAGTGG 720
661 AAGAGCAAGTTCAAAATATCATCAAGAGACATCAAGGCTTTAAAGAGGCAACAAAGTGG 720
721 ACAAGGTGGTTGATCTGTGAGCTGCCAACACAGAGAGGTACAGTAATTTGTGTGGGCT 780
721 ACAAGGTGGTTGATCTGTGAGCTGCCAACACAGAGAGGTACAGTAATTTGTGTGGGCT 780
781 TTAATGACACCATGAGAAATCTTGTGGCTGTGTGAGACAGAAATGAGGTGAGATTTCTC 840
781 TTAATGACACCATGAGAAATCTTGTGGCTGTGTGAGACAGAAATGAGGTGAGATTTCTC 840
841 CTTCCACCTGTGATGCATGCTGTGTGTATGAGAAATGTTCTTTCATTAATGAGAGCC 900
841 CTTCCACCTGTGATGCATGCTGTGTGTATGAGAAATGTTCTTTCATTAATGAGAGCC 900
901 CTGAGAAACATTTTGTACACGAGGCTGATTTGTCATCCGACGAGAACACTTTGATTTG 960
901 CTGAGAAACATTTTGTACACGAGGCTGATTTGTCATCCGACGAGAACACTTTGATTTG 960
961 GTGAGATGACCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGTTGATTTCTTG 1020
961 GTGAGATGACCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGTTGATTTCTTG 1020
1021 TGGGGGCTGTGATCAAGCAACATCTATGATGATTAACAACATCTGGGAAACAAATGATG 1080
1021 TGGGGGCTGTGATCAAGCAACATCTATGATGATTAACAACATCTGGGAAACAAATGATG 1080
1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGGAAATCTCCAGAGCAACGTTG 1140

1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGGAAATCTCCAGAGCAACGTTG 1140
1141 TTGATGATATGATCAACAGCAATGCCATCTCTATGAGCCGTGGAACATCCAGACATG 1200
1141 TTGATGATATGATCAACAGCAATGCCATCTCTATGAGCCGTGGAACATCCAGACATG 1200
1201 TTGTTGTTATTAAGTATGCTTACGTAGGAGGACAGCAAGAGCCATGATGATGACA 1260
1201 TTGTTGTTATTAAGTATGCTTACGTAGGAGGACAGCAAGAGCCATGATGATGACA 1260
1261 CTTGAGATATTTCAATGGGTGGGAAAGACACCATTTGTTTGCACAAACATGCCAGATT 1320
1261 CTTGAGATATTTCAATGGGTGGGAAAGACACCATTTGTTTGCACAAACATGCCAGATT 1320
1321 CCGCTTATGCTGCTCTTATCTTGAACCTTGTGCTTCTTGTGAGCTCAGACTAGAA 1380
1321 CCGCTTATGCTGCTCTTATCTTGAACCTTGTGCTTCTTGTGAGCTCAGACTAGAA 1380
1381 TCGATTTAAAGCTGAATAATGAGGAAATTCACCTCATTCACCAAGTTGCTACATCC 1440
1381 TCGATTTAAAGCTGAATAATGAGGAAATTCACCTCATTCACCAAGTTGCTACATCC 1440
1441 TCGACTATTCACCAAGGCTCTCTGTGTTCCACCGGTTACACAGTGTGATGATGAT 1500
1441 TCGACTATTCACCAAGGCTCTCTGTGTTCCACCGGTTACACAGTGTGATGATGAT 1500
1501 CAAGGACGCTGCAATGCTGGAACAAATGAGAGGCTGTGTTGATTTGAGCCCGAGAGA 1560
1501 CAAGGACGCTGCAATGCTGGAACAAATGAGAGGCTGTGTTGATTTGAGCCCGAGAGA 1560
1561 ATTAATGATTTCTGAGATGAAGTGAAGCATGAGGACCGAAGAAATATTAATTTGGGGTAG 1620
1561 ATTAATGATTTCTGAGATGAAGTGAAGCATGAGGACCGAAGAAATATTAATTTGGGGTAG 1620
1621 CTTAGCTGAATGTTTATGTTAATATATGTTTGTCTTATTAATTTTGCAGAGTATTTGAA 1680
1621 CTTAGCTGAATGTTTATGTTAATATATGTTTGTCTTATTAATTTTGCAGAGTATTTGAA 1680
1681 TGCATCAGCTTCAATTAATGCTTTAGAGCGGGGCAATTCGTTTACTGAGAACTGAATG 1740
1681 TGCATCAGCTTCAATTAATGCTTTAGAGCGGGGCAATTCGTTTACTGAGAACTGAATG 1740
1741 AATGATATTAATTTTGTGT 1760
1741 AATGATATTAATTTTGTGT 1760

RESULT 2
US-10-718-952-1
; Sequence 1, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-1

Query Match 100.0%; Score 1760; DB 19; Length 1760;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCTTTATTCCTTTTGTATTTTCAATTCATTCCTTAATCTTTGTGAAAAATATGTTCA 60
DB 1 CTCTCTTTATTCCTTTTGTATTTTCAATTCATTCCTTAATCTTTGTGAAAAATATGTTCA 60
QY 61 TCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACGAGACTGAGATTCAGTCCG 120
DB 61 TCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACGAGACTGAGATTCAGTCCG 120
QY 121 TGTACAACTACGAAACCAACCGAATCTTGTACAGAAACAGAAATGGCACTTACAGTGA 180
DB 121 TGTACAACTACGAAACCAACCGAATCTTGTACAGAAACAGAAATGGCACTTACAGTGA 180
QY 181 TTGTCAAAACCAATCTGTCAAAATAGCAATTTAAAAACAATCATCTAATGTTCTAAATTA 240
DB 181 TTGTCAAAACCAATCTGTCAAAATAGCAATTTAAAAACAATCATCTAATGTTCTAAATTA 240
QY 241 GGGTAAATGCTTGTGGGTTGGGTTGAAACAAAGGCTCAACCTCACCGGTGTGTTATTG 300
DB 241 GGGTAAATGCTTGTGGGTTGGGTTGAAACAAAGGCTCAACCTCACCGGTGTGTTATTG 300
QY 301 CTAAACGAGAGGCAATTTCAATGGGCTACAAAGCAAGATTTCAACAAAGCCAAATTAATCT 360
DB 301 CTAAACGAGAGGCAATTTCAATGGGCTACAAAGCAAGATTTCAACAAAGCCAAATTAATCT 360
QY 361 GCTCCCTCAACCAAGGCTCAGTATCCGAGTTGGTCTTTCAGAGGAGAGAAATCTAATG 420
DB 361 GCTCCCTCAACCAAGGCTCAGTATCCGAGTTGGTCTTTCAGAGGAGAGAAATCTAATG 420
QY 421 CCCCATTTCAAGAGCTGTCTTCAATGTTTAACTTGAACAATTTGTTGGGGAATGG 480
DB 421 CCCCATTTCAAGAGCTGTCTTCAATGTTTAACTTGAACAATTTGTTGGGGAATGG 480
QY 481 AATATCGAACAATGAACTGTGCTGATGCCATGGCCAGGGCAAAAGTGTGTAATGATG 540
DB 481 AATATCGAACAATGAACTGTGCTGATGCCATGGCCAGGGCAAAAGTGTGTAATGATG 540
QY 541 TCGAAGAGAGTTGAGGCTTACATGGAATTCATGCTTCACTCCCGGAATCTATGACC 600
DB 541 TCGAAGAGAGTTGAGGCTTACATGGAATTCATGCTTCACTCCCGGAATCTATGACC 600
QY 601 CGGATTTCAATGCTGCCAACCAAGAGAGGCTGCCAACACGTCATCAAGGGCAAAAGC 660
DB 601 CGGATTTCAATGCTGCCAACCAAGAGAGGCTGCCAACACGTCATCAAGGGCAAAAGC 660
QY 661 AAGAGCAAGTTCAAAATCATCAAGACATCAAGGCTTTAAGAGAGCCAAAGTGG 720
DB 661 AAGAGCAAGTTCAAAATCATCAAGACATCAAGGCTTTAAGAGAGCCAAAGTGG 720
QY 721 ACAAGGTGGTTGACTGTGGAAGTCCCAACAGAGAGTACAGTAATTTGTTGGGGCC 780
DB 721 ACAAGGTGGTTGACTGTGGAAGTCCCAACAGAGAGTACAGTAATTTGTTGGGGCC 780
QY 781 TTAAATGACACCAATGAGAAATCTTGTGCTGTGTGACAGAAATGAGGCTGAGATTTC 840
DB 781 TTAAATGACACCAATGAGAAATCTTGTGCTGTGTGACAGAAATGAGGCTGAGATTTC 840
QY 841 CTTCACCTTTGATGTCATTTGCTGTGTATGAAAAATGTTCTTTCATTAATGAGAGCC 900
DB 841 CTTCACCTTTGATGTCATTTGCTGTGTATGAAAAATGTTCTTTCATTAATGAGAGCC 900
QY 901 CTCAAGAACATTTTGTACAGAGGCTGATGATCTTCCATCGGAGAGAACATTTGATG 960
DB 901 CTCAAGAACATTTTGTACAGAGGCTGATGATCTTCCATCGGAGAGAACATTTGATG 960
QY 961 GTGAGAGATGATTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTTG 1020
DB 961 GTGAGAGATGATTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTTG 1020

QY 1021 TGGGGCTGGTATCAAGCCCAACATCTATAGTGTACAAACCATCTGGGAAACAATGATG 1080
DB 1021 TGGGGCTGGTATCAAGCCCAACATCTATAGTGTACAAACCATCTGGGAAACAATGATG 1080
QY 1081 GATGATCTTTTCGGCTCCACAACCTTTCGTTCCAAGGAATCTCCAAAGAGCAAGTTG 1140
DB 1081 GATGATCTTTTCGGCTCCACAACCTTTCGTTCCAAGGAATCTCCAAAGAGCAAGTTG 1140
QY 1141 TTGATGATATGCTCAACAGCATATGCTCTATAGAGCTGTGAAACATCCAGACCATG 1200
DB 1141 TTGATGATATGCTCAACAGCATATGCTCTATAGAGCTGTGAAACATCCAGACCATG 1200
QY 1201 TTGTGTTATTAAGTATGTCCTTACGTAGGGGACAGAAAGGCAATGAGATGATCA 1260
DB 1201 TTGTGTTATTAAGTATGTCCTTACGTAGGGGACAGAAAGGCAATGAGATGATCA 1260
QY 1261 CTTCAGAGATATTCATGAGGTGGAAGAGCAATTTTTCACAAACATGCGAGGATTT 1320
DB 1261 CTTCAGAGATATTCATGAGGTGGAAGAGCAATTTTTCACAAACATGCGAGGATTT 1320
QY 1321 CCTCTTAAGCTGCTCTATTAATCTTGAATGAGTGTCTTGTGAGCTCAGACTAGAA 1380
DB 1321 CCTCTTAAGCTGCTCTATTAATCTTGAATGAGTGTCTTGTGAGCTCAGACTAGAA 1380
QY 1381 TCGAGTTTAAAGCTGAAATGAGGGGAAATTCACATTCACCCAGTTGCTAACATCC 1440
DB 1381 TCGAGTTTAAAGCTGAAATGAGGGGAAATTCACATTCACCCAGTTGCTAACATCC 1440
QY 1441 TCACTACCTCAACCAAGGCTCCTCTGTTTCAACCGGGTACACAGAGTGTGAATGATG 1500
DB 1441 TCACTACCTCAACCAAGGCTCCTCTGTTTCAACCGGGTACACAGAGTGTGAATGATG 1500
QY 1501 CAAAGAGGCTGCAATGCTGGAAGAAACATATAGAGGCTTGTGTGATTTGCCCAAGA 1560
DB 1501 CAAAGAGGCTGCAATGCTGGAAGAAACATATAGAGGCTTGTGTGATTTGCCCAAGA 1560
QY 1561 AATACATGATTTCTCGAGTCAAGTGAAGATGGAACCGAAGAAATATATGTTGGGTG 1620
DB 1561 AATACATGATTTCTCGAGTCAAGTGAAGATGGAACCGAAGAAATATATGTTGGGTG 1620
QY 1621 CCTAGCTGAATGTTTATTTATTTATATATATGTTTGTCTTAATTTTCCAGATGTA 1680
DB 1621 CCTAGCTGAATGTTTATTTATTTATATATATGTTTGTCTTAATTTTCCAGATGTA 1680
QY 1681 TGCATCAGCTTCAATTAATCTTTAAGCGGGCATATCTGTTTACTAGAAACATGAATG 1740
DB 1681 TGCATCAGCTTCAATTAATCTTTAAGCGGGCATATCTGTTTACTAGAAACATGAATG 1740
QY 1741 AATGATGATTAATTTTGTCT 1760
DB 1741 AATGATGATTAATTTTGTCT 1760

RESULT 3
US-10-424-598-70167
; Sequence 70167, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5323) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 70167
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1
US-10-424-559-70167

Query Match 87.4%; Score 1538; DB 18; Length 1989;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
1 CTCTCTTATTCCTTTGTAATTCATTCATTCCTTAATCTTTGTAATAATTAATGTTCA 60
164 CTCTCTTATTCCTTTGTAATTCATTCATTCCTTAATCTTTGTAATAATTAATGTTCA 223
61 TCGAAGATTTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGACTGAGATTCAGTCCG 120
224 TCGAAGATTTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGACTGAGATTCAGTCCG 283
121 TGTAACTAAGAAAACCAAGAACTTTGTCAGAGAAAGAAATGGCACTTACAGTGA 180
284 TGTAACTAAGAAAACCAAGAACTTTGTCAGAGAAAGAAATGGCACTTACAGTGA 343
181 TGTCAAAACCAAAATCTGTCAAAATAGAAATTTAAACCAACATCAATGTTCTTAATTAG 240
344 TGTCAAAACCAAAATCTGTCAAAATAGAAATTTAAACCAACATCAATGTTCTTAATTAG 403
241 GGGTAATGCTTGGGTT- GGGGTGAAACAAAGGCTCAACCTTCAACGGTGTATT 299
404 GGGTAATGCTTGGGTTGGGGGTGGAACAAAGGCTCAACCTTCAACGGTGTATT 463
300 GGTAAACGAGAGGGGCTTTCATGGGCTACAAAGGACAAAGTTCACAAAGCAATTAATT 359
464 GGTAAACGAGAGGGGCTTTCATGGGCTACAAAGGACAAAGTTCACAAAGCAATTAATT 523
360 GGGTCCCTCACCACCAAGCTCAGCTACGAGTGGGTCTTCCAGGAGAGGAAATCTAT 419
524 GGGTCCCTCACCACCAAGCTCAGCTACGAGTGGGTCTTCCAGGAGAGGAAATCTAT 583
420 GGGTCCCTCACCACCAAGCTCAGCTACGAGTGGGTCTTCCAGGAGAGGAAATCTAT 479
584 GGGTCCCTCACCACCAAGCTCAGCTACGAGTGGGTCTTCCAGGAGAGGAAATCTAT 643
480 GATATCAGCAATGAACTGGCTGATGTCATGGGCGGAGGAAAGGTTTGAACATCGAT 539
644 GATATCAGCAATGAACTGGCTGATGTCATGGGCGGAGGAAAGGTTTGAACATCGAT 703
540 TTGCAAGAGCAGTTGAGGCTTACATGGAATCCATGCTTCCACTCCCGGAAATCTATGAC 599
704 TTGCAAGAGCAGTTGAGGCTTACATGGAATCCATGCTTCCACTCCCGGAAATCTATGAC 763
600 CCGGATTTATGCTGCTGCCAACCAAGAGCGTGGCCAAAGTCTCAACGGGCAACAAG 659
764 CCGGATTTATGCTGCTGCCAACCAAGAGCGTGGCCAAAGTCTCAACGGGCAACAAG 823
660 CAAGAGCAAGTTCACCAATCATCAAGAGCATCAAGAGGCTTTAAGGAAGCAACAAGTG 719
824 CAAGAGCAAGTTCACCAATCATCAAGAGCATCAAGAGGCTTTAAGGAAGCAACAAGTG 883
720 GACAAAGTGGTGTACTGTGACTGCTGCCAACAGAGAGGTAAGTAATTTGTTGGGCT 779
884 GACAAAGTGGTGTACTGTGACTGCTGCCAACAGAGAGGTAAGTAATTTGTTGGGCT 943
780 CTTAATGACACATGAGAAATCTCTTGCTGCTGTGAGACAGAAATGAGGCTGAGTTCT 839
944 CTTAATGACACATGAGAAATCTCTTGCTGCTGTGAGACAGAAATGAGGCTGAGTTCT 1003
840 CCTTCCACCTGTAAGGCAATGCTGTGTGTAAGGAAAAGTTCTTTCAATTAAGGAAGC 899
1004 CCTTCCACCTGTAAGGCAATGCTGTGTGTAAGGAAAAGTTCTTTCAATTAAGGAAGC 1063
900 CCTCAGAAACATTTTGTACAGGGCTGATGATCTTGCCATGCGGAGAACTTTGATT 959
1064 CCTCAGAAACATTTTGTACAGGGCTGATGATCTTGCCATGCGGAGAACTTTGATT 1123
960 GGTGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTGTTGATTCTT 1019
```

```
1124 GGTGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTGTTGATTCTT 1183
1020 GTGGGGGCTGGTATCAAGCCCAACATCTATAGTCAAGTTCAACCAATCTGGGAAACAATGAT 1079
1184 GTGGGGGCTGGTATCAAGCCCAACATCTATAGTCAAGTTCAACCAATCTGGGAAACAATGAT 1243
1080 GGTATGAATCTTTGGGCTCCAAACCTTCCGTTCCAAAGAAATCTCCAAAGCAAGTT 1139
1244 GGTATGAATCTTTGGGCTCCAAACCTTCCGTTCCAAAGAAATCTCCAAAGCAAGTT 1303
1140 GTTGAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
1304 GTTGAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1363
1200 GTTGTGTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
1364 GTTGTGTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1423
1424 ACTCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1483
1260 ACTCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319
1320 TCCCTCTAGCTGCTCCATTAATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1379
1484 TCCCTCTAGCTGCTCCATTAATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1543
1380 ATCAGATTTAAAGCTGAAATGAGGGAATTCACATGATTCACACCAAGTTGCTACATC 1439
1544 ATCAGATTTAAAGCTGAAATGAGGGAATTCACATGATTCACACCAAGTTGCTACATC 1603
1440 CTGAGCTACCTCACAAGGCTCTCTGTGTTTCAACGGGTACACAGTGTGATGATGATGATGATGAT 1499
1604 CTGAGCTACCTCACAAGGCTCTCTGTGTTTCAACGGGTACACAGTGTGATGATGATGATGATGAT 1663
1500 TCAAGAGCAGGCTGCAATCTGGAACAAATTAAGAGGCTGTGTTGATGATGATGATGATGATGATGAT 1559
1664 TCAAGAGCAGGCTGCAATCTGGAACAAATTAAGAGGCTGTGTTGATGATGATGATGATGATGATGAT 1723
1560 AATAACATGATTTCTGAGATACAGTGAAGATGAGGACGAAGAAATATATGTTGGGCTA 1619
1724 AATAACATGATTTCTGAGATACAGTGAAGATGAGGACGAAGAAATATATGTTGGGCTA 1783
1620 GCTTACGTAATGTTTATGTTAATATATGTTGCTTAAATTTTGAAGTATGATGATGATGATGATGATGAT 1679
1784 GCTTACGTAATGTTTATGTTAATATATGTTGCTTAAATTTTGAAGTATGATGATGATGATGATGATGAT 1843
1680 ATGATCAGCTTCAATTAATGCTTGAAGAGGAGGCAATTTGTTACTAGGAACATGATGAT 1739
1844 ATGATCAGCTTCAATTAATGCTTGAAGAGGAGGCAATTTGTTACTAGGAACATGATGAT 1903
1740 GAATGATATATTTTGTGT 1760
1904 GAATGATATATTTTGTGT 1924
```

```
RESULT 4
US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US2003074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
```

NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 9
 LENGTH: 1533
 TYPE: DNA
 ORGANISM: Glycine max
 US-10-025-003-9

Query Match 87.1%; Score 1533; DB 14; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

54 ATGTTGATCGAGAAATTTAAAGTTGAGTGTCTAATGTGAATGACCCGAGCTGAGATT 113
 1 AGTTTATCGAGAAATTTAAAGTTGAGTGTCTAATGTGAATGACCCGAGCTGAGATT 60
 114 CAGTCCGTGTACAACTACGAAACCAACGAACTTGTTCACGAGAACGAAATGGCACTAT 173
 61 CAGTCCGTGTACAACTACGAAACCAACGAACTTGTTCACGAGAACGAAATGGCACTAT 120
 174 CAGTGTATGTCAAAACCCCAATCTGTCAAAATGAAATTTAAACCAACATCCATGTTCT 233
 121 CAGTGTATGTCAAAACCCCAATCTGTCAAAATGAAATTTAAACCAACATCCATGTTCT 180
 234 AAATTAGGGGTAAATGCTGTGGGTGGGTGAAACAAAGGCTCAACCCCTCACGGGTGT 293
 181 AAATTAGGGGTAAATGCTGTGGGTGGGTGAAACAAAGGCTCAACCCCTCACGGGTGT 240
 294 GTTATGTCTAACCGAAGGGCAATTTCAATGGCTTACAAAGAACAAAGATTCAACAGCCAT 353
 241 GTTATGTCTAACCGAAGGGCAATTTCAATGGCTTACAAAGAACAAAGATTCAACAGCCAT 300
 354 TACTTTGGCTCCCTCACCCCAAGCTTACGATTCGAGTTGGTCTTCCAGGAGAGAA 413
 301 TACTTTGGCTCCCTCACCCCAAGCTTACGATTCGAGTTGGTCTTCCAGGAGAGAA 360
 414 ATCTATGCCCATCTTAAGAGCTGCTTCCAAATGGTTAACCTGACGAACTGTGTTGGG 473
 361 ATCTATGCCCATCTTAAGAGCTGCTTCCAAATGGTTAACCTGACGAACTGTGTTGGG 420
 474 GGAATGGATATACGAACTGAACTGTGATGCAATGGCCCAAGGCAAGGTTGAC 533
 421 GGAATGGATATACGAACTGAACTGTGATGCAATGGCCCAAGGCAAGGTTGAC 480
 534 ATCGATTTGCAAGAGCAATTTGAGGCTTACATGAAATCCATGCTTCACTCCCGGAATC 593
 481 ATCGATTTGCAAGAGCAATTTGAGGCTTACATGAAATCCATGCTTCACTCCCGGAATC 540
 594 TATGACCCGGATTTCAATGCTGCTCAACCAAGAGAGGTCCAACCAAGCTCATCAAGGCG 653
 541 TATGACCCGGATTTCAATGCTGCTCAACCAAGAGAGGTCCAACCAAGCTCATCAAGGCG 600
 654 ACAAAAGCAAGAGCAATTTCAACCAATCATCAAGAGAGGCTTTTAAAGAAAGCAC 713
 601 ACAAAAGCAAGAGCAATTTCAACCAATCATCAAGAGAGGCTTTTAAAGAAAGCAC 660
 714 AAAGTGAACAAGTGTGTGTAATCTGTGTAATCTGTGTAATCTGTGTAATCTGTGTAAT 773
 661 AAAGTGAACAAGTGTGTGTAATCTGTGTAATCTGTGTAATCTGTGTAATCTGTGTAAT 720
 774 GTGGGCTTAATGACCAATGAGAAATCTGTTGGCTGTGTAATCTGTGTAATCTGTGTAAT 833
 721 GTGGGCTTAATGACCAATGAGAAATCTGTTGGCTGTGTAATCTGTGTAATCTGTGTAAT 780
 834 ATTTCTCTTCAACCTTGTATGCAATGCTGTGTGTATGAAATGCTTCTTCAATTAAT 893
 781 ATTTCTCTTCAACCTTGTATGCAATGCTGTGTGTATGAAATGCTTCTTCAATTAAT 840
 894 GGAAGCCCTCAAGAACTTTTGTACAGAGGCTGATGATCTTCCATCGGAGAAACACT 953
 841 GGAAGCCCTCAAGAACTTTTGTACAGAGGCTGATGATCTTCCATCGGAGAAACACT 900
 954 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGAT 1013

DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
 1014 TTCTTTGTGGGGCTGTATCAAGCCCAATCATATAGTCAATCAACCATCTGGAAAC 1073
 961 TTCTTTGTGGGGCTGTATCAAGCCCAATCATATAGTCAATCAACCATCTGGAAAC 1020
 1074 AATGATGTATGAATCTTTTGGCTCCCAAACTTTTCCGTTCCAAAGAAATCTTCAAGAGC 1133
 1021 AATGATGTATGAATCTTTTGGCTCCCAAACTTTTCCGTTCCAAAGAAATCTTCAAGAGC 1080
 1134 AACGTTGTGATGATATGCTCAACAGCAATGCAATCTCTATGAGGCTGGTGAATCCA 1193
 1081 AACGTTGTGATGATATGCTCAACAGCAATGCAATCTCTATGAGGCTGGTGAATCCA 1140
 1194 GACCATGTTGTTATTAATGATGCTTACGTAAGGGGACAGCAAGAGCCATGAT 1253
 1141 GACCATGTTGTTATTAATGATGCTTACGTAAGGGGACAGCAAGAGCCATGAT 1200
 1254 GAGTACATTTCAAGATATTTCAATGGGTGGAAGAGCAATGTTTTCACACATGC 1313
 1201 GAGTACATTTCAAGATATTTCAATGGGTGGAAGAGCAATGTTTTCACACATGC 1260
 1314 GAGATTTCCCTTGTAGCTGCTCTATATCTTGAATGATGCTGCTGAGCTCAGC 1373
 1261 GAGATTTCCCTTGTAGCTGCTCTATATCTTGAATGATGCTGCTGAGCTCAGC 1320
 1374 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTCACATCTTCCACCCAGTTGCT 1433
 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTCACATCTTCCACCCAGTTGCT 1380
 1434 ACCATCTCAAGTACTCTCAACCAAGGCTCTCTGTTTCCACCGGTACACAGTGTGAT 1493
 1381 ACCATCTCAAGTACTCTCAACCAAGGCTCTCTGTTTCCACCGGTACACAGTGTGAT 1440
 1494 GCATTTGCAAGAGGTCATATGCTGAAATCATTAAGGCTGTGTGATTTGGCC 1553
 1441 GCATTTGCAAGAGGTCATATGCTGAAATCATTAAGGCTGTGTGATTTGGCC 1500
 1554 CCAGAGATTAATGATATTTCTCGATGACAGTGA 1586
 1501 CCAGAGATTAATGATATTTCTCGATGACAGTGA 1533

RESULT 5
 US-10-718-952-9
 ; Sequence 9, Application US/10718952
 ; Publication No. US20040128713A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hitz, William
 ; APPLICANT: Sebastian, Scott
 ; APPLICANT: Grace, John
 ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
 ; TITLE OF INVENTION: SACCCHARIDES AND PHYTIC ACID
 ; FILE REFERENCE: BB-1077-C
 ; CURRENT APPLICATION NUMBER: US/10/718,952
 ; PRIOR FILING DATE: 2003-11-21
 ; PRIOR APPLICATION NUMBER: 08/835,751
 ; PRIOR FILING DATE: APRIL 8, 1997
 ; PRIOR APPLICATION NUMBER: PCT/US98/06822
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 9
 ; LENGTH: 1533
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-718-952-9
 Query Match 87.1%; Score 1533; DB 19; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	5	TTGTTCAATCGAAGATTTTAAAGTTGAGTGTCTAAATGGAAGTACACCGAATCTGAATT	113
Db	1	ATGTTCAATCGAAGATTTTAAAGTTGAGTGTCTAAATGGAAGTACACCGAATCTGAAGTT	60
OY	114	CAGTCCGTTGTCACATACGAAGACCAACCACTTGTTCAAGAAACAGGAATGGACCTAT	173
Db	61	CAGTCCGTTGTCACATACGAAGACCAACCACTTGTTCAAGAAACAGGAATGGACCTAT	120
OY	174	CAGTGAATTTGTCAAAACCCCAATCTGTCAAAATACGAATTTAAACCAATCATGTTTCT	233
Db	121	CAGTGAATTTGTCAAAACCCCAATCTGTCAAAATACGAATTTAAACCAATCATGTTTCT	180
OY	234	AAATTAGGGGTAATGCTTTGTGGGTTGGGGTGGAAACAACGGCTCAACCCCTACCGTGTG	293
Db	181	AAATTAGGGGTAATGCTTTGTGGGTTGGGGTGGAAACAACGGCTCAACCCCTACCGTGTG	240
OY	294	GTTATTTGCTAACCGAGAGGGCATTTTCAATGGGCTACAAAGACACAGATTCAACACCCAT	353
Db	241	GTTATTTGCTAACCGAGAGGGCATTTTCAATGGGCTACAAAGAGCAAGATTCAACACCCAT	300
OY	354	TACTTTGGCTCCCTACCCCAAGCCCTACATATCCGATGGGGTCCCTACAGAGAGAGAA	413
Db	301	TACTTTGGCTCCCTACCCCAAGCCCTACATATCCGATGGGGTCCCTACAGAGAGAGAA	360
OY	414	ATCTATGCCCCATTCAMAGGCTGTCTTCAAATGTTAAACCTGACGATTTGTGTTGGG	473
Db	361	ATCTATGCCCCATTCAMAGGCTGTCTTCAAATGTTAAACCTGACGATTTGTGTTGGG	420
OY	474	GGATGGGATATCAGAACATGAACCTGGCTGATGCAATGCGCAGGGGCAAAAGTGTGGAC	533
Db	421	GGATGGGATATCAGAACATGAACCTGGCTGATGCAATGCGCAGGGGCAAAAGTGTGGAC	480
OY	534	ATCGATTTGCAGAAAGCAGTTGAGGCTCTTACATGAGATCATGCTTCCACTCCCGAATC	593
Db	481	ATCGATTTGCAGAAAGCAGTTGAGGCTCTTACATGAGATCATGCTTCCACTCCCGAATC	540
OY	594	TATGACCCCGAATTTTATGCTGTGCCACAACAGAGAGCGTCCACACATCATCAAGGGC	653
Db	541	TATGACCCCGAATTTTATGCTGTGCCACAACAGAGAGCGTCCACACATCATCAAGGGC	600
OY	654	ACAAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGGGCTTTAAAGAAAGCCAC	713
Db	601	ACAAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGGGCTTTAAAGAAAGCCAC	660
OY	714	AAAGTGAACAAGTGTGTGACTGTGACTGTGCACTGCAACACAGAGAGGTACAGTAATTTGGTT	773
Db	661	AAAGTGAACAAGTGTGTGACTGTGACTGTGCACTGCAACACAGAGAGGTACAGTAATTTGGTT	720
OY	774	GTGGGCTTATATGACACATGAGAAATCTCTTGGTGTGTGAGACAGAAATGAGCGGTAG	833
Db	721	GTGGGCTTATATGACACATGAGAAATCTCTTGGTGTGTGAGACAGAAATGAGCGGTAG	780
OY	834	ATTTCTCCTTCACCTTGATGCACTTGCTGTGTGTTATGAAATGTTCTTTCAATTAAT	893
Db	781	ATTTCTCCTTCACCTTGATGCACTTGCTGTGTGTTATGAAATGTTCTTTCAATTAAT	840
OY	894	GGAAGCCTTCAGAACACTTTTGTACCAAGGGCTGATTAATCTTGCCATGCGAGAACCT	953
Db	841	GGAAGCCTTCAGAACACTTTTGTATCAAGGGCTGATTAATCTTGCCATGCGAGAACCT	900
OY	954	TTGATTTGGTGAAGATCTTCAAGAGTGTGAGACCAAAATGAATCTGTGTGGTTGAT	1013
Db	901	TTGATTTGGTGAAGATCTTCAAGAGTGTGAGACCAAAATGAATCTGTGTGGTTGAT	960
OY	1014	TTCTTTGTGGGGCTGTGATCAAGCCAACTCTATATAGTCAGTTACACCATCTGGGAAAC	1073
Db	961	TTCTTTGTGGGGCTGTGATCAAGCCAACTCTATATAGTCAGTTACACCATCTGGGAAAC	1020
OY	1074	AATGATGATTAAGAACTTTGGCTTCCAAACTTTCCGTTCCAGAGAAATCTCCAAAGAC	1133
Db	1021	AATGATGATTAAGAACTTTGGCTTCCAAACTTTCCGTTCCAGAGAAATCTCCAAAGAC	1080

OY	113	AACTTTGTAATGATATGATGCAAGAAATGCATCTCTATGAGCTGGGAAATCCA	1193
Db	1081	AACCTTTGATGATATGATGCAAGCAATGCCATCTCTATGAGCTGGTGAATCCA	1140
OY	1194	GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAGGGACAGCAAGAGCCATGAT	1253
Db	1141	GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAGGGACAGCAAGAGCCATGAT	1200
OY	1254	GAGTACACTTCAGAGATATTCATGAGGTGGGAAAGACCAATGTTTGCAACACATGC	1313
Db	1201	GAGTACACTTCAGAGATATTCATGAGGTGGGAAAGACCAATGTTTGCAACACATGC	1260
OY	1314	GAGGATTCCTCTTAGCTGCTCCTATATCTTTGACTTGTGCTTCTTGTGCTGAC	1373
Db	1261	GAGGATTCCTCTTAGCTGCTCCTATATCTTTGACTTGTGCTTCTTGTGCTGAC	1320
OY	1374	ACTGTAAATCGAGTTTAAAGCTGAAATAGAGGAAAAATTCACATCATTCACCCAGTTGCT	1433
Db	1321	ACTGTAAATCGAGTTTAAAGCTGAAATAGAGGAAAAATTCACATCATTCACCCAGTTGCT	1380
OY	1434	ACCATCCCTACGCTACCTACCAAGGCTCCTCTGGTTTCCAACGGGTACACCAAGTGTGAAT	1493
Db	1381	ACCATCCCTACGCTACCTACCAAGGCTCCTCTGGTTTCCAACGGGTACACCAAGTGTGAAT	1440
OY	1494	GCATTGTCAAGACGCGTCAGATCTGTGAAAACATAATGAGGGCTTGTGTGATTTGCC	1553
Db	1441	GCATTGTCAAGACGCGTCAGATCTGTGAAAACATAATGAGGGCTTGTGTGATTTGCC	1500
OY	1554	CCAGAGAAATACATGATTTCTGAGATTAAGAAGA	1586
Db	1501	CCAGAGAAATACATGATTTCTGAGATTAAGAAGA	1533

```

RESULT 6
US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025.003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-5

```

	Query Match	Similarity	84.2% Best Local	Score 1482; Pred. No. of Matches 1532	DB 14; Mismatches 1;	Length 1533; Indels 0;	Gaps 0;
QY	54	ATGTTTCATCGGAATTTTAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATT	113				
DB	1	ATGTTTCATCGGAATTTTAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATT	60				
QY	114	CAGTCCGTTGTCACACTACGAAACACCGAACTTGTTCACGGAACACAGAAATGGCACTTAT	173				
DB	61	CAGTCCGTTGTCACACTACGAAACACCGAACTTGTTCACGGAACACAGAAATGGCACTTAT	120				
QY	174	CAGTGGATTTGCAAAACCAATCTGTCAATACGAATTTAAACCAATTCATGTTCTCT	233				

Db 121 CAGTGGATTGTCAAAACCAAACTGTCAATACGAATTTAAACCAACATCATGTTCCCT 180
Qy 234 AAATTAGGGGTATGCTTGGGTTGGGGTGGAAAAACAACGGCTCAACCTCACCGGTGGT 233
Db 181 AAATTAGGGGTATGCTTGGGTTGGGGTGGAAAAACAACGGCTCAACCTCACCGGTGGT 240
Qy 294 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGAGCAAGATTCAACAGCAAT 353
Db 241 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGAGCAAGATTCAACAGCAAT 300
Qy 354 TACTTGGCTCCCTCAACCAAGCTCAAGCTATCCAGTTGGGTCCTTCCAGGAGAGAA 413
Db 301 TACTTGGCTCCCTCAACCAAGCTCAAGCTATCCAGTTGGGTCCTTCCAGGAGAGAA 360
Qy 414 ATCTATGCCCCCATTCAGAGGCTGCTTCCAAATGTTAACTTCAACGACATTTGTTGGG 473
Db 361 ATCTATGCCCCCATTCAGAGGCTGCTTCCAAATGTTAACTTCAACGACATTTGTTGGG 420
Qy 474 GGATGGGATATCAGAACATGAACCTGGCTGATGGCCAGGGCAAGGTTGTTGAC 533
Db 421 GGATGGGATATCAGAACATGAACCTGGCTGATGGCCAGGGCAAGGTTGTTGAC 480
Qy 534 ATCGATTTCAGAAAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAAATC 593
Db 481 ATCGATTTCAGAAAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAAATC 540
Qy 594 TATGACCCCGGATTTCTTGTCTGCCAACAAAGAGAGCTGCCAAACGTCATCAAGGGC 653
Db 541 TATGACCCCGGATTTCTTGTCTGCCAACAAAGAGAGCTGCCAAACGTCATCAAGGGC 600
Qy 654 ACAAGCAAGAGCAAGTTCAACAAATCATCAAAAGATCAAGGCGTTTAAAGAAAGCAC 713
Db 601 ACAAGCAAGAGCAAGTTCAACAAATCATCAAAAGATCAAGGCGTTTAAAGAAAGCAC 660
Qy 714 AAAGTGAACAAGGTGTTGTAAGTGTGAGTCTGCCACACAGAGAGTACAGTAATTTGGT 773
Db 661 AAAGTGAACAAGGTGTTGTAAGTGTGAGTCTGCCACACAGAGAGTACAGTAATTTGGT 720
Qy 774 GTGGGCTTAATGACACATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
Db 721 GTGGGCTTAATGACACATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
Qy 834 ATTTCTCCTTCCACCTTGTATGCAATGCTGTGTATGGAATAATGTTCCCTTCAATTAAT 893
Db 781 ATTTCTCCTTCCACCTTGTATGCAATGCTGTGTATGGAATAATGTTCCCTTCAATTAAT 840
Qy 894 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATGATCTTGGCAATCGGAGAAACAT 953
Db 841 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATGATCTTGGCAATCGGAGAAACAT 900
Qy 954 TTGATTGCTGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGTTGAT 1013
Db 901 TTGATTGCTGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGTTGAT 960
Qy 1014 TTCCCTGTGGGGGCTGTATCAAGCAACATCTATGTCAAGTTACAAACATCGGGAAAC 1073
Db 961 TTCCCTGTGGGGGCTGTATCAAGCAACATCTATGTCAAGTTACAAACATCGGGAAAC 1020
Qy 1074 AATGATGATATGATCTTCCGCTCCCAAACTTTCCTTCCAAAGAAAATCTTCCAGAGC 1133
Db 1021 AATGATGATATGATCTTCCGCTCCCAAACTTTCCTTCCAAAGAAAATCTTCCAGAGC 1080
Qy 1134 AACGTTGTTGATATGCTCAACAGCAATGCAATCTCTATGAGCTGTGTAACATCCA 1193
Db 1081 AACGTTGTTGATATGCTCAACAGCAATGCAATCTCTATGAGCTGTGTAACATCCA 1140
Qy 1194 GACCATGTTGTTGTTAATGATGCTTACGTAGGGGAGCAGAAAGAGCAATGAT 1253
Db 1141 GACCATGTTGTTGTTAATGATGCTTACGTAGGGGAGCAGAAAGAGCAATGAT 1200
Qy 1254 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGCAATGTTTGGACAACATGCTC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGCAATGTTTGGACAACATGCTC 1260

Qy 1314 GAGGATTCCTCTAGCTGCTCTTATCTTATCTTGAATGAGTCTTCTGTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTAGCTGCTCTTATCTTATCTTGAATGAGTCTTCTGTGAGCTCAGC 1320
Qy 1374 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTTCCACTATTCACCAAGTTGGT 1433
Db 1321 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTTCCACTATTCACCAAGTTGGT 1380
Qy 1434 ACCATCTCAGCTACTTCAACCAAGCTCTCTGTGTTCCACCGGATACCAAGTGTGAAT 1493
Db 1381 ACCATCTCAGCTACTTCAACCAAGCTCTCTGTGTTCCACCGGATACCAAGTGTGAAT 1440
Qy 1494 GCATTGTCAACAGCTGCTCAATGCTGAAATCATTAATGAGGCTTGTGATTTGGCC 1553
Db 1441 GCATTGTCAACAGCTGCTCAATGCTGAAATCATTAATGAGGCTTGTGATTTGGCC 1500
Qy 1554 CCAAGAAATACATGATTTCTCGAGTCAAGTGA 1586
Db 1501 CCAAGAAATACATGATTTCTCGAGTCAAGTGA 1533

RESULT 7
US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-718-952-5

Query Match 84.2%; Score 1482; DB 19; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 54 ANGTTCATCGAAGATTTTAAAGTTGAGTGTCCAAATGTGAATACACCGAGCTGAGATT 113
Db 1 ANGTTCATCGAAGATTTTAAAGTTGAGTGTCCAAATGTGAATACACCGAGCTGAGATT 60
Qy 114 CAGTCCGTGTACACTACGAACCAACCGAATTTTTCAGAGAAACAGAAATGGCACCTAT 173
Db 61 CAGTCCGTGTACACTACGAACCAACCGAATTTTTCAGAGAAACAGAAATGGCACCTAT 120
Qy 174 CAGTGGATGTGAACCCAAATCTGTCAAAATGCAATTTTAAACCAACATCCATGTTCCCT 233
Db 121 CAGTGGATGTGAACCCAAATCTGTCAAAATGCAATTTTAAACCAACATCCATGTTCCCT 180
Qy 234 AAATTAGGGGTATGCTTGGGTTGGGGTGGAAAAACAACGGCTCAACCTCACCGGTGGT 293
Db 181 AAATTAGGGGTATGCTTGGGTTGGGGTGGAAAAACAACGGCTCAACCTCACCGGTGGT 240
Qy 294 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGAGCAAGATTCAACAGCAAT 353
Db 241 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGAGCAAGATTCAACAGCAAT 300
Qy 354 TACTTGGCTCCCTCAACCAAGCTCAAGCTATCCAGTTGGGTCCTTCCAGGAGAGAA 413


```
|||||
Db 301 TACTTGGCTCCCTCAACCCAGGCTCAGCTATCCGAGTTGGGCTTTCAGAGGAGAGAA 360
Qy 414 ATCTATGCCCCATGCAAGAGCCCTGCTTCCAAATGGTTAAACCTGACGACATTTGTTGGG 473
Db 361 ATCTATGCCCCATGCAAGAGCCCTGCTTCCAAATGGTTAAACCTGACGACATTTGTTGGG 420
Qy 474 GGATGGGATATGACAAATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 533
Db 421 GGATGGGATATGACAAATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
Qy 534 ATCGATTTTGCAGAGCAGTTGAGGCCCTTACATGGAATCCATGCTTCCACTCCCGGAATC 593
Db 481 ATCGATTTTGCAGAGCAGTTGAGGCCCTTACATGGAATCCATGCTTCCACTCCCGGAATC 540
Qy 594 TATGACCCCGGATTTCTATTGCTGCAACCAAGAGAGCGTCCCAACAAGCTCATCAAGGAC 653
Db 541 TATGACCCCGGATTTCTATTGCTGCAACCAAGAGAGCGTCCCAACAAGCTCATCAAGGAC 600
Qy 654 ACAAGCAGAGCAGATTTCAACAAATCATCAAGACATCAAGGCGTTTAAAGAGGACACC 713
Db 601 ACAAGCAGAGCAGATTTCAACAAATCATCAAGACATCAAGGCGTTTAAAGAGGACACC 660
Qy 714 AAAGTGGACAGGTTGTTGTACTGTGACTGCCAACAGAGAGGTACAGTAATTTGGTT 773
Db 661 AAAGTGGACAGGTTGTTGTACTGTGACTGCCAACAGAGAGGTACAGTAATTTGGTT 720
Qy 774 GTGGGCTTAATGACCAATGAGAAATCTCTGGCTGCTGTGACAGAAATAGAGGCTGAG 833
Db 721 GTGGGCTTAATGACCAATGAGAAATCTCTGGCTGCTGTGACAGAAATAGAGGCTGAG 780
Qy 834 ATTTCTCTTCCACCTTGTATGCGCATTTGTTGTTATGGAATATGTTCTTCTTCAATTAAT 893
Db 781 ATTTCTCTTCCACCTTGTATGCGCATTTGTTGTTATGGAATATGTTCTTCTTCAATTAAT 840
Qy 894 GGAAGCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCATCGGAGAAACCT 953
Db 841 GGAAGCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCATCGGAGAAACCT 900
Qy 954 TTGATTTGGTGAAGATCACTTCAAGAGTGGTCAAGCCAAATATGAAATCTGTGTTGTTGAT 1013
Db 901 TTGATTTGGTGAAGATCACTTCAAGAGTGGTCAAGCCAAATATGAAATCTGTGTTGTTGAT 960
Qy 1014 TTCTCTTGGGGGCTGTATCAAGCCAAATCTATGTCAGTTTCAACCATCTGGGAAAC 1073
Db 961 TTCTCTTGGGGGCTGTATCAAGCCAAATCTATGTCAGTTTCAACCATCTGGGAAAC 1020
Qy 1074 AATGATGATGATGAATCTTTCGGCTCCACAACTTTCCGTTCCAAAGGAAATCTCCAGAGC 1133
Db 1021 AATGATGATGATGAATCTTTCGGCTCCACAACTTTCCGTTCCAAAGGAAATCTCCAGAGC 1080
Qy 1134 AACGTTTGGTGAATATGTCACACAGCATATGCTCTCTATGAGCGCTGTGAAATCTCA 1193
Db 1081 AACGTTTGGTGAATATGTCACACAGCATATGCTCTCTATGAGCGCTGTGAAATCTCA 1140
Qy 1194 GACCATGTTGTTTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGGACATGAT 1253
Db 1141 GACCATGTTGTTTATTAAGTATGTCCTTACGTAGGGGACAGCAATGAGGACATGAT 1200
Qy 1254 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACCATTTGTTTGCACAAACATGC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACCATTTGTTTGCACAAACATGC 1260
Qy 1314 GAGGATTCCTCTTGAAGCTCTCTATTAATCTTGGAATTTGGTCTTCTGTAGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTTGAAGCTCTCTATTAATCTTGGAATTTGGTCTTCTGTAGAGCTCAGC 1320
Qy 1374 ACTAGATATGAGTTTAAAGCTGAATGAGGAAATTCACATCTTCCACCAGTTGGCT 1433
Db 1321 ACTAGATATGAGTTTAAAGCTGAATGAGGAAATTCACATCTTCCACCAGTTGGCT 1380
Qy 1434 ACCATCTCAGTACCTCAACAAAGCTCTCTGTTTCCACGGGTAACACAGTGTGTAAT 1493
|||||
```

```
Db 1381 ACCATCTCAGTACCTCAACAAAGGCTCTCTGTTTCCACGGGTACCAAGTGTGTAAT 1440
Qy 1494 GCATTGTCAAAAGCAGCGTGCATATGCTGAAAAATATATGAGGGCTTGTGATTTGGCC 1553
Db 1441 GCATTGTCAAAAGCAGCGTGCATATGCTGAAAAATATATGAGGGCTTGTGATTTGGCC 1500
Qy 1554 CCAGAAATTAACATGATTTCTCGAGTACAAAGTGA 1586
Db 1501 CCAGAAATTAACATGATTTCTCGAGTACAAAGTGA 1533

RESULT 8
US-10-424-599-70166
; Sequence 70166, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 70166
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34371C.1
US-10-424-599-70166

Query Match 16.8%; Score 295; DB 18; Length 594;
Best Local Similarity 100.0%; Pred. No. 1,1e-145;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1466 GGTTCCACCGGGTACACGAGTGTGAATGCAATTTGCAAGCAGCGTGAATGCTGGAAA 1525
Db 300 GGTTCCACCGGGTACACGAGTGTGAATGCAATTTGCAAGCAGCGTGAATGCTGGAAA 359
Qy 1526 CATATATGAGGGCTTGTGTTGATTTGGCCCAAGATTAACATGATTTCTGAGTACAACTG 1585
Db 360 CATATATGAGGGCTTGTGTTGATTTGGCCCAAGATTAACATGATTTCTGAGTACAACTG 419
Qy 1586 AAGCATGGGACCGAAGAAATATATGTTGGGGTACCTTACCTGAATGTTTATGTTAATA 1645
Db 420 AAGCATGGGACCGAAGAAATATATGTTGGGGTACCTTACCTGAATGTTTATGTTAATA 479
Qy 1646 ATATGTTGCTTATTAATTTTGAAGTATTAATGATGATGATGATGATGATGATGATGAT 1705
Db 480 ATATGTTGCTTATTAATTTTGAAGTATTAATGATGATGATGATGATGATGATGATGAT 539
Qy 1706 AGCGGGGCAATATTTCTGTTTACTAGGAACATGAATGAATGAATGAATTTGTGT 1760
Db 540 AGCGGGGCAATATTTCTGTTTACTAGGAACATGAATGAATGAATGAATTTGTGT 594

RESULT 9
US-10-424-599-116597
; Sequence 116597, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
```

NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 116597
LENGTH: 661
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(661)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_76298C.1
US-10-424-599-116597

Query Match 15.7%; Score 276; DB 18; Length 661;
Best Local Similarity 100.0%; Pred.No.1.5e-135;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 818 CAGAAATGAGGCTGATGATTTCTCTCCACCTTGATGACCATGCTTGATGAGAAA 877
Db 4 CAGAAATGAGGCTGATGATTTCTCTCCACCTTGATGACCATGCTTGATGAGAAA 63
Qy 878 TGTTCCTTCATTAATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATGATCTTGC 937
Db 64 TGTTCCTTCATTAATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATGATCTTGC 123
Qy 938 CATCGGAGAGAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 997
Db 124 CATCGGAGAGAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
Qy 998 ATCTGTGTTGATGATTTCTGTGGGGCTGGATCAAGCCAACTATAGTCAGTTA 1057
Db 184 ATCTGTGTTGATGATTTCTGTGGGGCTGGATCAAGCCAACTATAGTCAGTTA 243
Qy 1058 CAACCATCTGGGAGAACATGATGATGATGATCTTTC 1093
Db 244 CAACCATCTGGGAGAACATGATGATGATGATCTTTC 279

RESULT 10

US-10-424-599-22663
Sequence 22663, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 22663
LENGTH: 634
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_120467C.1
US-10-424-599-22663

Query Match 8.3%; Score 146; DB 18; Length 634;
Best Local Similarity 100.0%; Pred.No.2.2e-66;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 870 ATGAAAAATGTTCTTCATTAATGAGAGCCCTCAGAACCTTTTACAGGGCTGATT 929
Db 116 ATGAAAAATGTTCTTCATTAATGAGAGCCCTCAGAACCTTTTACAGGGCTGATT 175
Qy 930 GATCTTGCCATGCGAGAGAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
Db 176 GATCTTGCCATGCGAGAGAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 235

Qy 990 AAAATGAATCTGTGTGTTGATT 1015
Db 236 AAAATGAATCTGTGTGTTGATT 261

RESULT 11

US-10-025-003-11
Sequence 11, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 11
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-11

Query Match 8.3%; Score 146; DB 14; Length 1533;
Best Local Similarity 100.0%; Pred.No.2.3e-66;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 870 ATGAAAAATGTTCTTCATTAATGAGAGCCCTCAGAACCTTTTACAGGGCTGATT 929
Db 817 ATGAAAAATGTTCTTCATTAATGAGAGCCCTCAGAACCTTTTACAGGGCTGATT 876
Qy 930 GATCTTGCCATGCGAGAGAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
Db 877 GATCTTGCCATGCGAGAGAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 936
Qy 990 AAAATGAATCTGTGTGTTGATT 1015
Db 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 12

US-10-025-003-13
Sequence 13, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 13
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-13

Thu Jun 9 09:20:42 2005

us-10-718-952-1.011.rnpb

Page 11

Search completed: June 8, 2005, 19:34:52
Job time : 1093.38 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:47:55 ; Search time 92 Seconds
(without alignments)
2143.998 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MFIEFKVCECVKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq16dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	510	2	AAW79740 Soybean w
2	510	100.0	510	8	ADQ14499 Mutant so
3	510	100.0	510	8	ADQ14491 Wild type
4	510	100.0	510	8	ADG82000 Soybean m
5	510	100.0	510	2	ADG81994 Soybean m
6	409	80.2	510	2	AAW79741 Soybean m
7	409	80.2	510	8	ADQ14495 Mutant so
8	409	80.2	510	8	ADG81998 Soybean m
9	253	49.6	510	8	ADQ14503 Mutant so
10	253	49.6	510	8	ADQ14505 Wild type
11	253	49.6	510	8	ADG82004 Soybean m
12	253	49.6	510	8	ADG82006 Soybean m
13	236	46.3	510	8	ADQ14501 Mutant so
14	236	46.3	510	8	ADG82002 Soybean m
15	69	13.5	536	2	AAW79740 Soybean w
16	68	13.3	394	5	ABP51436 Human MDP
17	68	13.3	510	4	ABG47286 MIP synth
18	67	13.1	380	3	AAQ32503 Arabidops
19	67	13.1	446	3	AAQ32502 Arabidops
20	67	13.1	446	3	AAQ09862 Arabidops
21	67	13.1	505	8	ADG73738 Myo-inosi
22	67	13.1	510	2	AAW97882 Maize myo
23	67	13.1	510	3	AAQ50575 Arabidops
24	67	13.1	510	4	AAQ48935 Brassica
25	67	13.1	510	8	ADN73525 Thale cre

ALIGNMENTS

26	67	13.1	510	8	ADP43920 Rice L-my
27	67	13.1	511	3	AAQ09861 Arabidops
28	67	13.1	511	3	AAQ32501 Arabidops
29	67	13.1	534	3	AAQ09860 Arabidops
30	67	13.1	581	3	AAQ50574 Arabidops
31	67	13.1	645	3	AAQ50573 Arabidops
32	50	9.8	510	2	AAW96259 Phytate P
33	43	8.4	512	8	ADP43919 Porterei
34	31	6.1	98	3	AAQ44825 Zea mays
35	29	5.7	565	4	ABG5497 Drosophi1
36	20	3.9	45	3	AAQ50576 Arabidops
37	20	3.9	78	5	ABP07716 Human ORF
38	20	3.9	113	3	AAQ12422 Zea mays
39	20	3.9	138	3	AAQ19004 Zea mays
40	20	3.9	174	3	AAQ19003 Zea mays
41	19	3.7	128	3	AAQ32705 Zea mays
42	19	3.7	130	3	AAQ32704 Zea mays
43	19	3.7	153	3	AAQ32703 Zea mays
44	19	3.7	251	3	AAQ52875 Arabidops
45	19	3.7	316	3	AAQ52874 Arabidops

RESULT 1
AAW79740
ID AAW79740 standard; protein; 510 AA.
XX
XX AAW79740;
AC
XX
DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
XX soybean wild-type myo-inositol 1-phosphate synthase.
DE
XX
XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM
KM phytic acid.
XX
XX Glycine max; line LR13.
OS
XX
XX
PN WO9845448-A1.
XX
XX 15-OCT-1998.
PD
XX
XX 07-APR-1998; 98WO-US006822.
PF
XX
XX 08-APR-1997; 97US-00835751.
PR (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Hitz WD, Sebastian SA;
PI
XX
XX WPI: 1998-568353/48.
DR N-PSDB; AAW62440.
XX
XX Soybean plants containing altered myo-inositol 1-phosphate gene - useful for generating plants with altered levels of e.g. raffinose, stachyose, phytic acid, etc.
XX
XX Example 5; Page 45-47; 63pp; English.
PS
XX
XX This is the amino acid sequence of soybean myo-inositol 1-phosphate
CC
CC synthase (MI 1-PS) deduced from the coding region of an isolated cDNA
CC clone (see AAW62440) . MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. A mutant MI 1-PS (see AAW79741) has been
CC identified in soybean line LR33, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC in the LR33 gene sequence that resulted in a K396N substitution in the
CC mutant protein. The mutation results in a seed phenotype of very low
CC raffinose saccharide sugars, very high sucrose and low phytic acid. The
CC mutated nucleic acid is used to alter the raffinose saccharide, sucrose,
CC phytic acid and inorganic phosphate content of soybean seeds, leading to

CC useful soybean products, e.g. a seed phytic acid content of less than 17
CC ug/g, a seed content of raffinose and stachyose combined of less than
CC 14.5 ug/g, and a seed sucrose content greater than 200 ug/g. (Updated on
CC 17-OCT-2003 to standardise OS field)

XX Sequence 510 AA;

Query Match 100.0%; Score 510; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MFIEHFKECPNVKXTEIETISVNYETTELVEHNRNGTYQWIVKPSVKYEFTNTIHP 60
DB 1 MFIEHFKECPNVKXTEIETISVNYETTELVEHNRNGTYQWIVKPSVKYEFTNTIHP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDIOQANFGSLTQASAIRVGSFQGE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDIOQANFGSLTQASAIRVGSFQGE 120
QY 121 IYAPFKSLPWNPDIDIVFGGWDISNMNLADAMARAKYFDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLPWNPDIDIVFGGWDISNMNLADAMARAKYFDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDFIAANOEERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDNRNABEISPSLTVAIACMVNPVPIINSPONTFVPGJIDLAIAANT 300
DB 241 VGLNDTMENLLAAVDNRNABEISPSLTVAIACMVNPVPIINSPONTFVPGJIDLAIAANT 300
QY 301 LIIGDDFSGQTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNLSAPQTFRSKEISKS 360
DB 301 LIIGDDFSGQTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNLSAPQTFRSKEISKS 360
QY 361 NVVDNMNSNALIYEPGHPDHVVVIKYPYVGSKRAMEYTSIEIFMGKSTIVLHNTC 420
DB 361 NVVDNMNSNALIYEPGHPDHVVVIKYPYVGSKRAMEYTSIEIFMGKSTIVLHNTC 420
QY 421 EDSLILAAPIIIDVILVLAESTPIEFKANEKGKSHFHVATILSVLTAPLVPPTPIVN 480
DB 421 EDSLILAAPIIIDVILVLAESTPIEFKANEKGKSHFHVATILSVLTAPLVPPTPIVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 2

ID ADQ14499 standard; protein; 510 AA.
XX ADQ14499;

AC ADQ14499;

DT 23-SEP-2004 (first entry)

XX Mutant soybean myo-inositol 1-phosphate synthase polypeptide #2.

KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KW mutant; mutant.

XX Glycine max.
OS Synthetic.

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.

XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI; 2004-53135/51.

XX N-PSDB; ADQ14498.

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.

XX Example 8; SEQ ID NO 10; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents a mutant soybean myo-inositol
XX 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MFIEHFKECPNVKXTEIETISVNYETTELVEHNRNGTYQWIVKPSVKYEFTNTIHP 60
DB 1 MFIEHFKECPNVKXTEIETISVNYETTELVEHNRNGTYQWIVKPSVKYEFTNTIHP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDIOQANFGSLTQASAIRVGSFQGE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDIOQANFGSLTQASAIRVGSFQGE 120
QY 121 IYAPFKSLPWNPDIDIVFGGWDISNMNLADAMARAKYFDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLPWNPDIDIVFGGWDISNMNLADAMARAKYFDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDFIAANOEERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDNRNABEISPSLTVAIACMVNPVPIINSPONTFVPGJIDLAIAANT 300
DB 241 VGLNDTMENLLAAVDNRNABEISPSLTVAIACMVNPVPIINSPONTFVPGJIDLAIAANT 300
QY 301 LIIGDDFSGQTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNLSAPQTFRSKEISKS 360
DB 301 LIIGDDFSGQTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNLSAPQTFRSKEISKS 360
QY 361 NVVDNMNSNALIYEPGHPDHVVVIKYPYVGSKRAMEYTSIEIFMGKSTIVLHNTC 420
DB 361 NVVDNMNSNALIYEPGHPDHVVVIKYPYVGSKRAMEYTSIEIFMGKSTIVLHNTC 420
```


QY 421 EDSLLAAPITIDVLVLAELSTRIEFKAENEGKFSFHPVATITLSYLTAKAPLVPGTTPVN 480
DB 421 EDSLLAAPITIDVLVLAELSTRIEFKAENEGKFSFHPVATITLSYLTAKAPLVPGTTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 3
ADQ14491
ID ADQ14491 standard; protein; 510 AA.
AC ADQ14491;
DT 23-SEP-2004 (first entry)
DE Wild type soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX
KM Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
XX raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.
OS Glycine max.
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hltz WD, Sebastian SA, Grace DJ, Streit LG;
XX
PI WPI; 2004-533135/51.
DR N-PSDB; ADQ14490.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Claim 3, SEQ ID NO 2; 48bp; English.

The invention relates to a nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate synthase having a decreased capacity for the synthesis of myo-inositol 1-phosphate. The invention also relates to a chimeric gene operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant comprising the chimeric gene, a method of making the soybean plant, a seed of the soybean plant, a soy protein product derived from the processing of soybean seeds, a method of making or producing a soy protein product and a method of using a soybean plant homozygous for at least one gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate. The nucleic acid fragment and methods are useful for producing plants with decreased raffinose, stachyose and phytic acid content and increased sucrose and inorganic phosphate content, leading to valuable and useful soybean products. This sequence represents a wild type soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENFKYECNPNVKTETETETISVNYETTELVHNRNGTQMIYVKSVEKFEKTNINHP 60
DB 1 MEIENFKYECNPNVKTETETETISVNYETTELVHNRNGTQMIYVKSVEKFEKTNINHP 60
QY 61 KLGVMLVGMGNGNSTLTGVIYANREGISWATKDKIQOANTFGLTQASAIRVSGFOGEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIYANREGISWATKDKIQOANTFGLTQASAIRVSGFOGEE 120
QY 121 IYAPFKSLPMVNPDPDIFVGGMDISNNMLADAMARAKVFDIDLOKQAPYMESMLPLPGI 180
DB 121 IYAPFKSLPMVNPDPDIFVGGMDISNNMLADAMARAKVFDIDLOKQAPYMESMLPLPGI 180
QY 181 YDPPFIANOEERANNVVKGTKEOVQOIIIDIKAFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPPFIANOEERANNVVKGTKEOVQOIIIDIKAFKATKVDKVVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDNRNABISPSSTLVAIACWENVPIFNGSPONTFVGLIDLAIAANT 300
DB 241 VGLNDTMENLLAAVDNRNABISPSSTLVAIACWENVPIFNGSPONTFVGLIDLAIAANT 300
QY 301 LIGGDFKSGQTKKMSVAVDFLVGAGIKPTISIVSNHGNNDGMLSAPOFRSKEISKS 360
DB 301 LIGGDFKSGQTKKMSVAVDFLVGAGIKPTISIVSNHGNNDGMLSAPOFRSKEISKS 360
QY 361 NVVDDMVNSNAIILYEPGEHPHVVVYIKVPYVGDGSKRAMDEYTSRIFMGKSTTVLHNTC 420
DB 361 NVVDDMVNSNAIILYEPGEHPHVVVYIKVPYVGDGSKRAMDEYTSRIFMGKSTTVLHNTC 420
QY 421 EDSLLAAPITIDVLVLAELSTRIEFKAENEGKFSFHPVATITLSYLTAKAPLVPGTTPVN 480
DB 421 EDSLLAAPITIDVLVLAELSTRIEFKAENEGKFSFHPVATITLSYLTAKAPLVPGTTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 4
ADS82000
ID ADS82000 standard; protein; 510 AA.
XX
AC ADS82000;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase wild-type 3.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
OS Glycine max; line 29004UP01.
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
PI Hltz WD, Sebastian SA;
XX
PI WPI; 2004-639957/62.
DR N-PSDB; ADS81999.
XX

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX Example 8; SEQ ID NO 10; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences), where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC oligosaccharides in soy plants (and other legumes) can lead to flautence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

CC Sequence 510 AA;

Query March 100.0%; Score 510; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 510; Conservative 0; Indels 0;

QY 1 MFIENFKVCEPNVKYETEISGVNYETTELVEHNRNGTYQWIVPKSVKEFKTNHVP 60
DB 1 MFIENFKVCEPNVKYETEISGVNYETTELVEHNRNGTYQWIVPKSVKEFKTNHVP 60
QY 61 KLGVMLVWVGNGNSTLTGGVIANREGISWATKDKIQOANFGSLTQSAIRVSGFQEE 120
DB 61 KLGVMLVWVGNGNSTLTGGVIANREGISWATKDKIQOANFGSLTQSAIRVSGFQEE 120
QY 121 IYAPKSLLPWNPDPDIFVGGMDISNNMLADAMARAKYFPDIDLOQLAPYMSMLPLGI 180
DB 121 IYAPKSLLPWNPDPDIFVGGMDISNNMLADAMARAKYFPDIDLOQLAPYMSMLPLGI 180
QY 181 YDPDFIAANOEBERANNVIKGTQEQVOOIIKDIKAFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPDFIAANOEBERANNVIKGTQEQVOOIIKDIKAFKATKVDKVVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDRNEAIEISPTLVIAIACMENVPPTNGSPONTFVGLIDLAIANT 300
DB 241 VGLNDTMENLLAAVDRNEAIEISPTLVIAIACMENVPPTNGSPONTFVGLIDLAIANT 300
QY 301 LIIGGDFSGGOTKMSVVDPLVGAGIKPTSVSVNHHGNDGNNLSAPQFRSKIEIKS 360
DB 301 LIIGGDFSGGOTKMSVVDPLVGAGIKPTSVSVNHHGNDGNNLSAPQFRSKIEIKS 360
QY 361 NVVDQMNASNAILEPGEHPDHVVVIKVPYVGDSKRAMDEYTSIEFGKSTIVLANTC 420
DB 361 NVVDQMNASNAILEPGEHPDHVVVIKVPYVGDSKRAMDEYTSIEFGKSTIVLANTC 420
QY 421 EDSLIIADPIILDVLLAELSTRIEFKANEKGFHSFHPVATILSLVLTAAPLVPGTPVYN 480
DB 421 EDSLIIADPIILDVLLAELSTRIEFKANEKGFHSFHPVATILSLVLTAAPLVPGTPVYN 480

DB 421 EDSLIIADPIILDVLLAELSTRIEFKANEKGFHSFHPVATILSLVLTAAPLVPGTPVYN 480
QY 481 ALSKQAMLENIMRACVGLAPENNMIIEYK 510
DB 481 ALSKQAMLENIMRACVGLAPENNMIIEYK 510

RESULT 5
ADS81994
ID ADS81994 standard; protein: 510 AA.

AC ADS81994;

DT 18-NOV-2004 (first entry)

DE Soybean myo-inositol 1-phosphate synthase wild-type 1.

KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flautence.

OS Glycine max; cultivar Wye.

PN US2003074685-A1.

XX 17-APR-2003.

PF 11-MAR-2002; 2002US-00025003.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.

PI Hitz WD, Sebastian SA;

XX WPI: 2004-639957/62.

DR N-PSDB; ADS81993.

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX Claim 3; SEQ ID NO 2; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences), where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flautence

CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MFIENFKECNPVKVKTETETISVYNYETTELVHNRNGTYQMIVKPKSVKTEFKTNHVP 60
DB 1 MFIENFKECNPVKVKTETETISVYNYETTELVHNRNGTYQMIVKPKSVKTEFKTNHVP 60
QY 61 KLGWVLVGMGNGNSTLTGVIYANREGISMAWKDKIQOANYGSLTQASAIRVGSFOGEE 120
DB 61 KLGWVLVGMGNGNSTLTGVIYANREGISMAWKDKIQOANYGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPVNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKQALPYMESMLPLPGI 180
DB 121 IYAPFKSLIPVNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKQALPYMESMLPLPGI 180
QY 181 YDPDFIAANOEERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDFIAANOEERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDNRNAEISPSSTLYAIACMENVPPINGSPONTFVPGILDLAIANT 300
DB 241 VGLNDTMENLLAAVDNRNAEISPSSTLYAIACMENVPPINGSPONTFVPGILDLAIANT 300
QY 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGKMLSAPOFRSKEISKS 360
DB 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGKMLSAPOFRSKEISKS 360
QY 361 NVVDNMVNSNAIILEPGEHPDHVVVIXVYPVGDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDNMVNSNAIILEPGEHPDHVVVIXVYPVGDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLILAAPIIIDVLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPPTPVVN 480
DB 421 EDSLILAAPIIIDVLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPPTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILLEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILLEYK 510
```

RESULT 6
AAW79741
ID AAW79741 standard; protein; 510 AA.

XX AC AAW79741;
XX 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
DE Soybean mutant myo-inositol 1-phosphate synthase.
XX
KM Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW phytic acid.
XX
OS Glycine max; line LR33.
XX
PN MO9845448-A1.
XX
PD 15-OCT-1998.
XX
XX 07-APR-1998; 98MO-USO06822.
PF
XX 08-APR-1997; 97US-00835751.
PR
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA
XX Hiltz WD, Sebastian SA;
PI

XX MPI: 1998-568353/48.
DR N-PSDB; AAV62443.
XX

PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
XX phytic acid, etc.
XX

PS Example 5; Page 49-51; 63pp; English.

CC This is the amino acid sequence of a mutant soybean myo-inositol 1-
CC phosphate synthase (MI 1-PS) deduced from the coding region of an
CC isolated cDNA clone (see AAV62443). MI 1-PS is involved in glucose
CC metabolism to phytic acid, raffinose and stachyose. The MI 1-PS was
CC identified in soybean line LR33, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC in the LR33 gene sequence that resulted in a K396N substitution in the
CC mutant protein compared to wild-type MI 1-PS (see AAW79740). The mutation
CC results in a seed phenotype of very low raffinose saccharide sugars, very
CC high sucrose and low phytic acid. The mutated nucleic acid is used to
CC alter the raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds, leading to useful soybean products,
CC e.g. a seed phytic acid content of less than 17 ug/g, a seed content of
CC raffinose and stachyose combined of less than 14.5 ug/g, and a seed
CC sucrose content greater than 200 ug/g. (Updated on 17-Oct-2003 to
CC standardise OS field)

XX Sequence 510 AA;

Query Match 80.2%; Score 409; DB 2; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MFIENFKECNPVKVKTETETISVYNYETTELVHNRNGTYQMIVKPKSVKTEFKTNHVP 60
DB 1 MFIENFKECNPVKVKTETETISVYNYETTELVHNRNGTYQMIVKPKSVKTEFKTNHVP 60
QY 61 KLGWVLVGMGNGNSTLTGVIYANREGISMAWKDKIQOANYGSLTQASAIRVGSFOGEE 120
DB 61 KLGWVLVGMGNGNSTLTGVIYANREGISMAWKDKIQOANYGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPVNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKQALPYMESMLPLPGI 180
DB 121 IYAPFKSLIPVNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKQALPYMESMLPLPGI 180
QY 181 YDPDFIAANOEERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDFIAANOEERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDNRNAEISPSSTLYAIACMENVPPINGSPONTFVPGILDLAIANT 300
DB 241 VGLNDTMENLLAAVDNRNAEISPSSTLYAIACMENVPPINGSPONTFVPGILDLAIANT 300
QY 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGKMLSAPOFRSKEISKS 360
DB 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGKMLSAPOFRSKEISKS 360
QY 361 NVVDNMVNSNAIILEPGEHPDHVVVIXVYPVGDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDNMVNSNAIILEPGEHPDHVVVIXVYPVGDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLILAAPIIIDVLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPPTPVVN 480
DB 421 EDSLILAAPIIIDVLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPPTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILLEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILLEYK 510
```

RESULT 7
ADQ14495
ID ADQ14495 standard; protein; 510 AA.

XX ADQ14495;
XX
XX 23-SEP-2004 (first entry)
XX
XX Mutant soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX
XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
XX raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
XX mutant; mutein.
XX
XX Glycine max.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 396 /note= "Wild type Lys substituted by Asn"
XX
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX 26-APR-1999; 99US-00299315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-533135/51.
XX N-PSDB; ADQ14494.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Claim 9; SEQ ID NO 6; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents a mutant soybean myo-inositol
XX 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 510 AA;
XX
XX Query Match 80.2%; Score 409; DB 8; Length 510;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MFIEHFKECPNVKYTELEIQSVYNYETTELVHNNRGTYQWIVKPKSVKYEFTNTINHP 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1 MFIEHFKECPNVKYTELEIQSVYNYETTELVHNNRGTYQWIVKPKSVKYEFTNTINHP 60
Qy KLGVMVLVGMGNGNSTLTGVIANREGISWATKDKIQOANVFGSLTQASAIRVGSFOGEE 120
Db KLGVMVLVGMGNGNSTLTGVIANREGISWATKDKIQOANVFGSLTQASAIRVGSFOGEE 120
Qy IYAPFKSLLPVNPDDIVFGGWDISNNMLADAMARAVFDIQLQOLRPYMESMLPLPGI 180
Db IYAPFKSLLPVNPDDIVFGGWDISNNMLADAMARAVFDIQLQOLRPYMESMLPLPGI 180
Qy YDPDFIAANOEBERNANNVTKGTQEQVOQIITKDIKAFKEATQVDKVVVLTMTANTERYSNLV 240
Db YDPDFIAANOEBERNANNVTKGTQEQVOQIITKDIKAFKEATQVDKVVVLTMTANTERYSNLV 240
Qy VGLNDTMENLLAAVDRBEAEISPTLYAIACVMENVPFINGSPONTVPGLIDLAIRNT 300
Db VGLNDTMENLLAAVDRBEAEISPTLYAIACVMENVPFINGSPONTVPGLIDLAIRNT 300
Qy 301 LIGDDFKSGQTKKSVLVDFLVGAGIKPTISVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Db 301 LIGDDFKSGQTKKSVLVDFLVGAGIKPTISVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Qy 361 NVVDDMNNSNAIIXEPEBHPDVVVIKVPYVGDGSKRAMDEYSEIFMGKSTVLHNTC 420
Db 361 NVVDDMNNSNAIIXEPEBHPDVVVIKVPYVGDGSKRAMDEYSEIFMGKSTVLHNTC 420
Qy 421 EDSLAAPIIIDVLVLAELSTRIEFKANEKGKFPHPVATISLYLTKAPLVPPGTPVNV 480
Db 421 EDSLAAPIIIDVLVLAELSTRIEFKANEKGKFPHPVATISLYLTKAPLVPPGTPVNV 480
Qy 481 ALSKQRAMLENIMKACVGLAPENNMILEYK 510
Db 481 ALSKQRAMLENIMKACVGLAPENNMILEYK 510
RESULT 8
AD881998
ID AD881998 standard; protein; 510 AA.
XX
XX AD881998;
XX
XX 18-NOV-2004 (first entry)
XX
XX Soybean myo-inositol 1-phosphate synthase mutant #1.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
XX phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
XX flatulence; mutant.
XX
XX Glycine max; line LR33.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 396 /note= "Wild-type Lys substituted by Arg"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX WPI; 2004-639957/62.
XX N-PSDB; AD881997.
XX

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 9; SEQ ID NO 6; 34bp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a mutant myo-
CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 80.2%; Score 409; DB 8; Length 510;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFIEHFKCECPVKTETETISGVVYVETTELVEHNRNGTYQVIYKPKYKEFKNNIHP 60
DB 1 MFIEHFKCECPVKTETETISGVVYVETTELVEHNRNGTYQVIYKPKYKEFKNNIHP 60
QY 61 KLGVWLVMGNGNGSTLTGCVIANREGISMATKDIQOANYGSLTQASAIRVGSFOGEE 120
DB 61 KLGVWLVMGNGNGSTLTGCVIANREGISMATKDIQOANYGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPVNPDPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYEMSLPLPGI 180
DB 121 IYAPFKSLIPVNPDPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYEMSLPLPGI 180
QY 181 YDPPFIANOSERRANVTKGTQOEVOQIINDIKAFKATKDKVYVMTANTERYSNLV 240
DB 181 YDPPFIANOSERRANVTKGTQOEVOQIINDIKAFKATKDKVYVMTANTERYSNLV 240
QY 241 VGLNDYTMENLAAVDRNEAEISPTLVAIACWENVPFINSPOTFVPGILDILAIANT 300
DB 241 VGLNDYTMENLAAVDRNEAEISPTLVAIACWENVPFINSPOTFVPGILDILAIANT 300
QY 301 LIIGDDPKSGGTQKMSVLVDFLVGAGIKPTSIIVSNHLGNDGNMLSAPOFRSKEISKS 360
DB 301 LIIGDDPKSGGTQKMSVLVDFLVGAGIKPTSIIVSNHLGNDGNMLSAPOFRSKEISKS 360
QY 361 NVVDNMSNAILVPERGHPDHVVVYKVPYVGSDEKRAMDETSTISIFNGKSTIYLAHTC 420
DB 361 NVVDNMSNAILVPERGHPDHVVVYKVPYVGSDEKRAMDETSTISIFNGKSTIYLAHTC 420
QY 421 EDSLAAPIIIDLVLLAEISTRIEFKANEKGFSHFVPAITISLYLTAPLPVPGTPVNV 480
DB 421 EDSLAAPIIIDLVLLAEISTRIEFKANEKGFSHFVPAITISLYLTAPLPVPGTPVNV 480

DB 421 EDSLAAPIIIDLVLLAEISTRIEFKANEKGFSHFVPAITISLYLTAPLPVPGTPVNV 480

QY 481 ALSKORAMLENTMRACVGLAPENNMLEYK 510

DB 481 ALSKORAMLENTMRACVGLAPENNMLEYK 510

RESULT 9

ADQ14503

ID ADQ14503 standard; protein; 510 AA.

AC ADQ14503;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #4.

KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;

KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;

KX mutant; mutein.

OS Glycine max.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98RO-US006822.

PR 26-APR-1999; 99US-00293315.

PR 11-MAR-2002; 2002US-00025003.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

PS MPI: 2004-533135/51.

PS N-PSDB; ADQ14502.

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,

XX useful for producing plants with decreased raffinose, stachyose, and

XX phytic acid and increased sucrose, leading to valuable and useful soybean

XX products.

XX Example 8; SEQ ID NO 14; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-

XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate

XX synthase having a decreased capacity for the synthesis of myo-inositol 1-

XX phosphate. The invention also relates to a chimeric gene operably linked

XX to suitable regulatory sequences, where expression of the chimeric gene

XX results in a decrease in expression of an endogenous or native gene

XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant

XX comprising the chimeric gene, a method of making the soybean plant, a

XX seed of the soybean plant, a soy protein product derived from the

XX processing of soybean seeds, a method of making or producing a soy

XX protein product and a method of using a soybean plant homozygous for at

XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having

XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The

XX nucleic acid fragment and methods are useful for producing plants with

XX decreased raffinose, stachyose and phytic acid content and increased

XX sucrose and inorganic phosphate content, leading to valuable and useful

XX soybean products. This sequence represents a mutant soybean myo-inositol

XX 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match	49.6%;	Score 253;	DB 8;	Length 510;
Best Local Similarity	99.7%;	Pred. No. 1.6e-244;		
Matches 353;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0
Qy	58	HVPKLGWLVGKGNNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFO	117	
Db	58	HVPKLGWLVGKGNNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFO	117	
Qy	118	GEEIYAPFKSLIPVYNPDIVFGWDISNMNLADAMARAKVPFDILOKOLRPYMSMLPL	177	
Db	118	GEEIYAPFKSLIPVYNPDIVFGWDISNMNLADAMARAKVPFDILOKOLRPYMSMLPL	177	
Qy	178	PGIYPPDPIAANOERRANNVIGCTQEOVOQIHKIKAKKEATKDKVYVLTWANTERS	237	
Db	178	PGIYPPDPIAANOERRANNVIGCTQEOVOQIHKIKAKKEATKDKVYVLTWANTERS	237	
Qy	238	NLVVGLNPTMENTLAAVDNREAEISPTLYAIACVMEVPTFNGSPQNTFVPGILDIAIA	297	
Db	238	NLVVGLNPTMENTLAAVDNREAEISPTLYAIACVMEVPTFNGSPQNTFVPGILDIAIA	297	
Qy	298	RNTLIGDDPFKSGQTKMKSVLVDPLVGAGIKPTISVYNHGLGNDGMNL SAPQTFRSKEI	357	
Db	298	RNTLIGDDPFKSGQTKMKSVLVDPLVGAGIKPTISVYNHGLGNDGMNL SAPQTFRSKEI	357	
Qy	358	SKSNVVDMDVNSNALLIYEPGEHPDHVYVTKIYVPYVDSGRANDEYSEIFMGK	411	
Db	358	SKSNVVDMDVNSNALLIYEPGEHPDHVYVTKIYVPYVDSGRANDEYSEIFMGK	411	
RESULT 10				
ID	ADQ14505	standard; protein; 510 AA.		
XX	ADQ14505;			
XX	ADQ14505;			
DT	23-SEP-2004	(first entry)		
DE	wild type soybean myo-inositol 1-phosphate synthase polypeptide #2.			
KW	soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;			
KW	raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.			
XX	Glycine max.			
OS				
XX	US2004128713-A1.			
PN				
PD	01-JUL-2004.			
XX				
PF	21-NOV-2003; 2003US-00718952.			
XX				
PR	08-APR-1997; 97US-00835751.			
PR	07-APR-1998; 98WO-US006822.			
PR	26-APR-1999; 98US-00299315.			
PR	11-MAR-2002; 2002US-00025003.			
XX				
PA	(HITZ/) HITZ W D.			
PA	(SEBA/) SEBASTIAN S A.			
PA	(GRAC/) GRACE D J.			
PA	(STRE/) STREIT L G.			
XX				
P1	Hitz WD, Sebastian SA, Grace DJ, Streit LG;			
XX				
XX	WPI; 2004-533135/51.			
DR	N-PsDB; ADQ14504.			
XX				
PT	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,			
PT	useful for producing plants with decreased raffinose, stachyose, and			
PT	phytic acid and increased sucrose, leading to valuable and useful soybean			
XX	products.			
XX				
XX	Claim 3; SEQ ID NO 16; 48bp; English.			
XX				
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-			

CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC	synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC	phosphate. The invention also relates to a chimeric gene operably linked
CC	to suitable regulatory sequences, where expression of the chimeric gene
CC	results in a decrease in expression of an endogenous or native gene
CC	encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC	comprising the chimeric gene, a method of making the soybean plant, a
CC	seed of the soybean plant, a soy protein product derived from the
CC	processing of soybean seeds, a method of making or producing a soy
CC	protein product and a method of using a soybean plant homozygous for at
CC	least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC	decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC	nucleic acid fragment and methods are useful for producing plants with
CC	decreased raffinose, stachyose and phytic acid content and increased
CC	sucrose and inorganic phosphate content, leading to valuable and useful
CC	soybean products. This sequence represents a wild type soybean myo-
CC	inositol 1-phosphate synthase polypeptide of the invention.
XX	
SQ	Sequence 510 AA:
Query Match	49.6%; Score 253; DB 8; Length 510;
Best Local Similarity	99.7%; Pred. No. 1.6e-244;
Matches 353; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
DY	58 HVPKLGWLVGMGNNSTLTGGVIARREGISMATKXKIQOANFGSLTQASAIRVGSFQ 117
Db	58 HVPKLGWLVGMGNNSTLTGGVIARREGISMATKXKIQOANFGSLTQASAIRVGSFQ 117
OY	118 GEETIAPPEKSLIPWNPPDDIVFGGMDISNMNLADAMARAKEYFDIDLQKQLRPYMESMLPL 177
Db	118 GEETIAPPEKSLIPWNPPDDIVFGGMDISNMNLADAMARAKEYFDIDLQKQLRPYMESWPL 177
OY	178 PGIVDPPEFIANOBERANNVIKGTROBOVOGIIDIKAFKEATKYDKRVVLMTANTERS 237
Db	178 PGIVDPPEFIANOBERANNVIKGTROBOVOGIIDIKAFKEATKYDKRVVLMTANTERS 237
OY	238 NLVVGNDTMENLLAAVDNRNAEISEPSTLYAIACMENVPPIINSPONTFPBGILDLAIA 297
Db	238 NLVVGNDTMENLLAAVDNRNAEISEPSTLYAIACMENVPPIINSPONTFPBGILDLAIA 297
OY	298 RNTLIIIGDDPFSGOTKKKSVLVDPLVAGAGIKPTSIIVSYNHLGNNDGNMLSAPOTFRSKEI 357
Db	298 RNTLIIIGDDPFSGOTKKKSVLVDPLVAGAGIKPTSIIVSYNHLGNNDGNMLSAPOTFRSKEI 357
OY	358 SKSNVVDDMVNSNALIYEPEHPDHVVVIKXVPYVGDSKRAMEDEYTSEIFMGSK 411
Db	358 SKSNVVDDMVNSNALIYEPEHPDHVVVIKXVPYVGDSKRAMEDEYTSEIFMGSK 411
RESULT 11	
ADS82004	
ID	ADS82004 standard; protein; 510 AA.
XX	
AC	ADS82004;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Soybean myo-inositol 1-phosphate synthase wild-type 4.
XX	
KM	Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KW	raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX	
OS	Glycine max; line 29018JP03.
XX	
PN	US2003074685-A1.
XX	
PD	17-APR-2003.
XX	
PF	11-MAR-2002; 2002US-00025003.
XX	
PR	08-APR-1997; 97US-00835751.
PR	07-APR-1998; 98WO-US006822.
XX	

PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PI Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
DR N-PSDB; ADS82003.
XX
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.

Example 8; SEQ ID NO 14; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 49.6%; Score 253; DB 8; Length 510;
Best Local Similarity 99.7%; Pred No. 1.6e-244;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HVPKLGWLVGVGNGNGSTLTGVIANREGISWATKDKIQANVFGSLTQSAIRVGSFQ 117
DB 58 HVPKLGWLVGVGNGNGSTLTGVIANREGISWATKDKIQANVFGSLTQSAIRVGSFQ 117
QY 118 GEEIYAPFKSLLPMPNPDIYVGGWDISNMNLADAMARAKVFDIDLOKQLRPMESMPL 177
DB 118 GEEIYAPFKSLLPMPNPDIYVGGWDISNMNLADAMARAKVFDIDLOKQLRPMESMPL 177
QY 178 PGIDPDPITANQGERANNVIKGTQEQVOQIIKIKAFKATKVDKVVLTANTERSY 237
DB 178 PGIDPDPITANQGERANNVIKGTQEQVOQIIKIKAFKATKVDKVVLTANTERSY 237
QY 238 NLVVGANDTMENLLAADVRENEAISPSTLYAIAQWENVPIFGSPONTFVGLIDLAIA 297
DB 238 NLVVGANDTMENLLAADVRENEAISPSTLYAIAQWENVPIFGSPONTFVGLIDLAIA 297
QY 298 RNTLIGSDPFSGGQTKMSVLVDPLVAGIKPTSTIVSNHLGNNGMLSAQOTRSKEI 357
DB 298 RNTLIGSDPFSGGQTKMSVLVDPLVAGIKPTSTIVSNHLGNNGMLSAQOTRSKEI 357
QY 358 SKSNVVDMMVNSNALIPEGEHPDHVVVIKYVPYVGDSCRAMDYTSIFMGCK 411
DB 358 SKSNVVDMMVNSNALIPEGEHPDHVVVIKYVPYVGDSCRAMDYTSIFMGCK 411

DB 358 SKSNVVDMMVNSNALIPEGEHPDHVVVIKYVPYVGDSCRAMDYTSIFMGCK 411

RESULT 12
ADSR82006
ID ADS82006 standard; protein; 510 AA.
XX
XX ADS82006;
AC
XX 18-NOV-2004 (first entry)
DT
XX
XX Soybean myo-inositol 1-phosphate synthase wild-type 2.
DE
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; cultivar Wye.
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
DR N-PSDB; ADS82005.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 3; SEQ ID NO 16; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

Sequence 510 AA;

Query Match 49.6%; Score 253; DB 8; Length 510;
 Best Local Similarity 99.7%; Pred. No. 1.6e-244;
 Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HVPKLGWLVWVGNGNSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIIVGSFQ 117
 DB 58 HVPKLGWLVWVGNGNSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIIVGSFQ 117
 QY 118 GEEIYAPKSLLPVNPDPDIVFGMDISNMNLADMAKAVPDDILOKQLPYMESMLPL 177
 DB 118 GEEIYAPKSLLPVNPDPDIVFGMDISNMNLADMAKAVPDDILOKQLPYMESMLPL 177
 QY 178 PGIVDPDFIAANOEBRANNVIKGTQEOVOQIIKDKAFKATKYDKVVLMTANTERSYS 237
 DB 178 PGIVDPDFIAANOEBRANNVIKGTQEOVOQIIKDKAFKATKYDKVVLMTANTERSYS 237
 QY 238 NLVVGINDTMENTLLAAVDRNEAISPSTLYAIACWENVFPINGSPOPTFVPGIDLAIA 297
 DB 238 NLVVGINDTMENTLLAAVDRNEAISPSTLYAIACWENVFPINGSPOPTFVPGIDLAIA 297
 QY 298 RNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSKEI 357
 DB 298 RNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSKEI 357
 QY 358 SKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
 DB 358 SKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411

RESULT 13
 ADQ14501
 ID ADQ14501 standard; protein; 510 AA.
 AC ADQ14501;
 XX 23-SEP-2004 (first entry)
 DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #3.
 XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
 KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
 KM mutant; mutein.
 XX Glycine max.
 OS Synthetic.
 XX US2004128713-A1.
 XX 01-JUL-2004.
 PD 21-NOV-2003; 2003US-00718952.
 PF 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-00299315.
 PR 11-MAR-2002; 2002US-00025003.
 XX (HTZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 XX WP1; 2004-533135/51.
 DR N-PSDB; ADQ14500.
 XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX

PS Claim 9; SEQ ID NO 12; 48bp; English.
 XX The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant, a
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents a mutant soybean myo-inositol
 CC 1-phosphate synthase polypeptide of the invention.
 XX

Sequence 510 AA:
 QY 176 PLPGIYDPDFIAANOEBRANNVIKGTQEOVOQIIKDKAFKATKYDKVVLMTANTER 235
 DB 176 PLPGIYDPDFIAANOEBRANNVIKGTQEOVOQIIKDKAFKATKYDKVVLMTANTER 235
 QY 236 YSNLVGINDTMENTLLAAVDRNEAISPSTLYAIACWENVFPINGSPOPTFVPGIDLA 295
 DB 236 YSNLVGINDTMENTLLAAVDRNEAISPSTLYAIACWENVFPINGSPOPTFVPGIDLA 295
 QY 296 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSK 355
 DB 296 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSK 355
 QY 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
 DB 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411

RESULT 14
 ADS82002
 ID ADS82002 standard; protein; 510 AA.
 XX ADS82002;
 AC 18-NOV-2004 (first entry)
 DT Soybean myo-inositol 1-phosphate synthase mutant #2.
 XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
 KW phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
 KM flatulence; mutant.
 XX Glycine max; line 29010CP01.
 OS Synthetic.
 OS Key Location/Qualifiers
 FH Misc-difference 87
 FT /note= "Wild-type Gly substituted by Asp"
 FT
 XX US2003074685-A1.
 FN 17-APR-2003.
 XX 11-MAR-2002; 2002US-00025003.
 XX 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: June 7, 2005, 16:58:47 ; Search time 24.5 Seconds
(without alignments)
2002.880 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MFIENKVCPCNVKKTETETI.....NIMRACVGLAPENMILEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR 79: *
2: p1r1: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	13.1	507	2 S52648	inositol-3-phospha
2	67	13.1	509	2 T08436	inositol-3-phospha
3	67	13.1	510	2 T04399	inositol-3-phospha
4	67	13.1	510	2 T50021	inositol-3-phospha
5	67	13.1	510	2 S60302	inositol-3-phospha
6	67	13.1	510	2 D84610	probable myo-inosl
7	67	13.1	511	2 T05017	inositol-3-phospha
8	67	13.1	512	2 T12438	inositol-3-phospha
9	50	9.8	510	2 T01647	inositol-3-phospha
10	48	9.4	511	2 T10964	inositol-3-phospha
11	15	2.9	430	2 T46317	hypothetical prote
12	11	2.5	525	2 T18569	inositol-3-phospha
13	13	2.5	555	2 A30902	inositol-3-phospha
14	12	2.4	520	2 S45452	inositol-3-phospha
15	8	1.6	210	2 S25544	minor outer capsid
16	8	1.6	378	2 P95844	nonstructural prot
17	8	1.6	429	2 E75107	probable dihydrol
18	8	1.6	429	2 AB2340	thiamin biosynthes
19	8	1.6	429	2 AB2340	hypothetical prote
20	7	1.4	42	2 T65746	tropomyosin - huma
21	7	1.4	96	1 WMVZP3	F3 protein - fowlp
22	7	1.4	97	2 JQ0527	capsid assembly pr
23	7	1.4	105	2 B82149	conserved hypotet
24	7	1.4	125	2 AB1995	hypothetical prote
25	7	1.4	147	2 T02273	hypothetical prote
26	7	1.4	157	2 S77352	hypothetical prote
27	7	1.4	178	2 S44910	ZK66.1 protein -
28	7	1.4	191	2 AH0431	probable exported
29	7	1.4	197	2 T15106	hypothetical prote

30	7	1.4	201	2 AH1812	hypothetical prote
31	7	1.4	209	2 A73313	probable v-type AT
32	7	1.4	212	2 T36864	probable cyclonexa
33	7	1.4	231	2 B83111	50S ribosomal prot
34	7	1.4	236	2 T33333	hypothetical prote
35	7	1.4	280	2 T47572	Machado-Joseph dis
36	7	1.4	282	2 AE2618	rRNA methylase (im
37	7	1.4	290	2 D97400	rRNA methylase (AP
38	7	1.4	294	2 C82497	ribose ABC transpo
39	7	1.4	294	2 AE2457	hypothetical prote
40	7	1.4	299	2 E84262	rhannosyl transfer
41	7	1.4	300	2 S36430	hypothetical prote
42	7	1.4	308	2 T06796	tropomyosin - huma
43	7	1.4	309	2 C83017	probable short cha
44	7	1.4	309	2 T02613	hypothetical prote
45	7	1.4	311	2 D64209	hydroxymethylgluta

ALIGNMENTS

RESULT 1

S52648 inositol-3-phosphate synthase (EC 5.5.1.4) - Citrus paradisi

C:Species: Citrus paradisi

C>Date: 19-Mar-1997 #sequence_rev1sion 09-May-1997 #text_change 09-Jul-2004

C:Accession: S52648

R:Holland, D. submitted to the EMBL Data Labratory, April 1994

A:Reference number: S52648

A:Accession: S52648

A:Molecule type: DNA

A:Residues: 1-507 <HOL>

A:Cross-references: UNIPROT:P42802; GB:Z32632; NID:G602564; PIDN:CAA83565.1; PID:G602565

C:Genetics:

A:Gene: INOI

C:Superfamily: myo-inositol-1-phosphate synthase

C:Keywords: Intramolecular lyase; isomerase

Query Match	13.1%; Score 67; DB 2; Length 507;
Best Local Similarity	100.0%; Pred. No. 1.6e-60;
Matches	67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	301 LIGGDDFKSGQTRKMSVLVDPLVGAGIKPTYSVYNHLGNNDGMNLSAQTRRSKISKS 360
DB	301 LIGGDDFKSGQTRKMSVLVDPLVGAGIKPTYSVYNHLGNNDGMNLSAQTRRSKISKS 360
QY	361 NVVDDMV 367
DB	361 NVVDDMV 367

RESULT 2

T08436 inositol-3-phosphate synthase (EC 5.5.1.4) [similarity] - rape

C:Species: Brassica napus (rape)

C>Date: 11-Jun-1999 #sequence_rev1sion 11-Jun-1999 #text_change 09-Jul-2004

C:Accession: T08436

R:Hussain, A.; Bourgeois, J.; Polvi, S.; Teang, E.; Keller, W.A.; George, F. submitted to the EMBL Data Library, August 1996

A:Reference number: Z16418

A:Accession: T08436

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-509 <HUS>

A:Cross-references: UNIPROT:Q96348; EMBL:U66307; NID:g1513227; PID:g1513228

C:Function:

A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-

C:Superfamily: myo-inositol-1-phosphate synthase

C:Keywords: intramolecular lyase; isomerase

Query Match	13.1%; Score 67; DB 2; Length 509;
Best Local Similarity	100.0%; Pred. No. 1.7e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
|||||
Db 300 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 359

QY 361 NVVDDMV 367
|||||
Db 360 NVVDDMV 366

RESULT 3
T04399
inositol-3-phosphate synthase (EC 5.5.1.4) - barley
C:Species: Hordeum vulgare (barley)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04399
R:Larson, S.R.; Raboy, V.
A:Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A:Reference number: Z14366
A:Accession: T04399
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-510 <LAR>
A:Cross-references: UNIPROT:O65195; EMBL:AF056325; NID:g3152730; PIDN:AAC17133.1; PID:g3
C:Genetics:
A:Gene: INO1
A:Map position: 4
C:Function:
A:Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inc
A:Pathway: inositol biosynthesis
A>Note: first step
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
|||||
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360

QY 361 NVVDDMV 367
|||||
Db 361 NVVDDMV 367

RESULT 4
T50021
inositol-3-phosphate synthase (EC 5.5.1.4) T3JP16.160 [similarity] - Arabidopsis thaliana
N:Alternate names: protein T3JP16.160
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50021
R:Bevan, M.; Zimmermann, W.; Gruenleisen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; Sm
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25027
A:Accession: T50021
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <BEV>
A:Cross-references: UNIPROT:Q9LX12; EMBL:AJ356332; GSPDB:GN00063; ATSP:T3JP16.160
A:Experimental source: cultivar Columbia; BAC clone T3JP16
C:Genetics:
A:Gene: ATSP:T3JP16.160
A:Map position: 5
A:Functions: 63/2; 86/2; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
|||||
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360

QY 361 NVVDDMV 367
|||||
Db 361 NVVDDMV 367

RESULT 5
S60302
inositol-3-phosphate synthase (EC 5.5.1.4) - Spirodela polyrrhiza
C:Species: Spirodela polyrrhiza
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60302
R:Smart, C.C.; Fleming, A.J.
A>Title: A plant gene with homology to D-myo-inositol-3-phosphate synthase is rapidly an
A:Reference number: S60302; MUID:94035182; PMID:8220483
A:Accession: S60302
A:Molecule type: mRNA
A:Residues: 1-510 <SMA>
A:Cross-references: UNIPROT:P42803; EMBL:Z11693; NID:g396209; PIDN:CAA7751.1; PID:g55864
C:Genetics:
A:Gene: cur1
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
|||||
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360

QY 361 NVVDDMV 367
|||||
Db 361 NVVDDMV 367

RESULT 6
D84610
probable myo-inositol 1-phosphate synthase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004
C:Accession: D84610
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; H
W.; Koo, H.; Molnar, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umeyam, L.; Tallon, L.
eaus, D.; Nierman, W.C.; White, C.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <STO>
A:Cross-references: UNIPROT:Q38862; GB:AE002093; NID:g4567202; PIDN:AAD33618.1; GSPDB:GNK
C:Genetics:
A:Gene: At2g2240
A:Map position: 2
C:Superfamily: Myo-inositol-1-phosphate synthase

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
|||||

Db 301 LIGDDFKSGQTKMKSVLVDPLVAGAIKPTISVYNHLGNNDGMNLSAPQTFRSKEISKS 360
QY 361 NVVDDMV 367
Db 361 NVVDDMV 367

RESULT 7

T05017
Inositol-3-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thaliana
N/Alternate names: protein T19P19.190
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T05017
R/Sevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hohnes, J.; Newes, H.W.;
submitted to the Protein Sequence Database, April 1998
A/Reference number: Z15394
A/Accession: T05017
A/Molecule type: DNA
A/Residues: 1-511 <REV>
A/Cross-references: UNIPROT:P42801; EMBL:AL022605
A/Experimental source: cultivar Columbia; BAC clone T19P19
C/Genetics:
A/Map position: 4
A/Intons: 64/2; 87/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3
A/Note: T19P19.190
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 13.1%; Score 67; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGDDFKSGQTKMKSVLVDPLVAGAIKPTISVYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 302 LIGDDFKSGQTKMKSVLVDPLVAGAIKPTISVYNHLGNNDGMNLSAPQTFRSKEISKS 361
QY 361 NVVDDMV 367
Db 362 NVVDDMV 368

RESULT 8

T12438
Inositol-3-phosphate synthase (EC 5.5.1.4) - common ice plant
C/Species: Mesembryanthemum crystallinum (common ice plant)
C/Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C/Accession: T12438
R/Ishitani, M.; Majumder, A.L.; Bornhauser, A.; Michalowski, C.B.; Jensen, R.G.; Bohner
Plant J. 9, 537-548, 1996
A/Title: Coordinate transcriptional induction of myo-inositol metabolism during enviro
A/Reference number: Z17518; M0ID:96208959; PMID:8624516
A/Accession: T12438
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-512 <ISH>
A/Cross-references: UNIPROT:Q40271; EMBL:U32511; NID:G975887; PID:AMB03687.1; PID:G9758
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase

Query Match 13.1%; Score 67; DB 2; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGDDFKSGQTKMKSVLVDPLVAGAIKPTISVYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 303 LIGDDFKSGQTKMKSVLVDPLVAGAIKPTISVYNHLGNNDGMNLSAPQTFRSKEISKS 362
QY 361 NVVDDMV 367
Db 363 NVVDDMV 369

RESULT 9

T01647
Inositol-3-phosphate synthase (EC 5.5.1.4) - maize
C/Species: Zea mays (maize)
C/Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01647
R/Jarson, S.R.; Raboy, V.
submitted to the EMBL Data Library, March 1998
A/Description: linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A/Reference number: Z14366
A/Accession: T01647
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-510 <LAR>
A/Cross-references: UNIPROT:Q9FPK7; EMBL:AF056326; NID:G3108052; PID:AACT5756.1; PID:G31
A/Experimental source: strain Early ACR; leaf
C/Genetics:
A/Gene: INO1
C/Function:
A/Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-ino
A/Pathway: inositol biosynthesis
A/Note: NAD cofactor
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 9.8%; Score 50; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 5.9e-43;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGDDFKSGQTKMKSVLVDPLVAGAIKPTISVYNHLGNNDGMNLSAPQ 350
Db 301 LIGDDFKSGQTKMKSVLVDPLVAGAIKPTISVYNHLGNNDGMNLSAPQ 350

RESULT 10

T10964
Inositol-3-phosphate synthase (EC 5.5.1.4) - kidney bean
N/Alternate names: 1L-myo-inositol 1-phosphate synthase
C/Species: Phaseolus vulgaris (kidney bean)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T10964
R/Wang, X.; Johnson, M.D.
submitted to the EMBL Data Library, October 1995
A/Reference number: Z17234
A/Accession: T10964
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-511 <MAN>
A/Cross-references: UNIPROT:Q41107; EMBL:U38920; NID:G1066282; PID:G1066283
A/Experimental source: strain Taylor's horticultural; root
C/Function:
A/Description: catalyzes reversible conversion of D-glucose 6-phosphate to 1L-myo-inositol
A/Pathway: myo-inositol biosynthesis
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 9.4%; Score 48; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 6.8e-41;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 DPLVAGAIKPTISVYNHLGNNDGMNLSAPQTFRSKEISKS NVVDDMV 367
Db 321 DPLVAGAIKPTISVYNHLGNNDGMNLSAPQTFRSKEISKS NVVDDMV 368

RESULT 11

T46317
Hypothetical protein DKFZp434A0612.1 - human
C/Species: Homo sapiens (man)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T46317
R/Duesterhoeft, A.; Lauber, J.; Newes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000

A:Reference number: 223035
A:Accession: T46317
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-430 <AAA>
A:Cross-references: UNIPROT:Q9NSU0; EMBL:AL137749
A:Experimental source: adult testis; clone DKF2p34A0612
C:Genetics:
A:Note: DKF2p34A0612.1
C:Superfamily: myo-inositol-1-phosphate synthase

Query Match 2.9%; Score 15; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 6.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 VIKKVPVYVGSKR 398
DB 253 VIKKVPVYVGSKR 267

RESULT 12

inositol-3-phosphate synthase (EC 5.5.1.4) - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18569; T20002
R:Ainscough, R.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z18979
A:Accession: T18569
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-525 <WIL>
A:Cross-references: UNIPROT:O18664; EMBL:AL033535; P1DN:CAA22132.1; CESP:VF13D12L.1
A:Experimental source: clone VF13D12L
R:Gajdasty, S.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19209
A:Accession: T20002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-525 <W12>
A:Cross-references: EMBL:269902; P1DN:CAA93771.1; GSPDB:GN00020; CESP:VF13D12L.1
A:Experimental source: clone C47D12
C:Genetics:
A:Gene: CESP:VF13D12L.1
A:Map position: 2
A:introns: 106/2; 287/1; 411/2
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 2.5%; Score 13; DB 2; Length 525;
Best Local Similarity 100.0%; Pred. No. 9.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 SIIVSYNHLGNNDG 343
DB 337 SIIVSYNHLGNNDG 349

RESULT 13

A30902
inositol-3-phosphate synthase (EC 5.5.1.4) [validated] - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein J0610; protein YJL153c
C:Species: *Saccharomyces cerevisiae*
C>Date: 18-Apr-1989 #sequence_revision 08-Sep-1995 #text_change 16-Aug-2004
C:Accession: S55160; B32209; S56935; S71644; A30827; A30902
R:Katsoulou, C.; Tzermita, M.; Alexandraki, D.
submitted to the EMBL Data Library, May 1995
A:Description: The complete sequence of a 40.7 kb segment located on the left arm of yeast hypochromosomal chromosome.
A:Reference number: S55159
A:Accession: S55160

A:Molecule type: DNA
A:Residues: 1-555 <KAT>
A:Cross-references: EMBL:X87371; NID:9854542; PID:9854544
R:Dean-Johnson, M.; Henry, S.A.
J. Biol. Chem. 264, 1274-1283, 1989
A:Title: Biosynthesis of inositol in yeast. Primary structure of myo-inositol-1-phosphate
A:Reference number: A32209; MUID:89093118; PMID:2642902
A:Accession: B32209
A:Molecule type: DNA
A:Residues: 23-35, 'RL', '37-81', 'FE', '83-87', 'TRNYAHVWR', '88', 'QQW', '92-103', 'WPRYWRISTWWS', '116-1
VISPQRLSFSFSAVL' <DEA>
A:Cross-references: EMBL:J04453
R:Katsoulou, C.; Tzermita, M.; Alexandraki, D.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56912
A:Accession: S56935
A:Molecule type: DNA
A:Residues: 1-555 <KAT>
A:Cross-references: EMBL:Z49428; NID:G1015570; PID:G1015571; MIPS:YJL153c
R:Katsoulou, C.; Tzermita, M.; Tavernarakis, N.; Alexandraki, D.
Yeast 12, 787-797, 1996
A:Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X
of chromosome XI.
A:Reference number: S71643; MUID:96408771; PMID:8613765
A:Accession: S71644
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-555 <KAF>
A:Cross-references: EMBL:X87371; NID:9854542; P1DN:CAA60802.1; PID:9854544
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C:Genetics:
A:Gene: SGD:INO1
A:Cross-references: SGD:S0003689; MIPS:YJL153c
A:Map position: 10L
C:Complex: homotetramer
C:Function: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-
A:pathway: inositol biosynthesis
A:Note: requires NAD
C:Superfamily: Myo-inositol-1-phosphate synthase
C:Keywords: cytosol; homotetramer; intramolecular lyase; isomerase; NAD

Query Match 2.5%; Score 13; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 INGSPONTFVPGT 291
DB 315 INGSPONTFVPGT 327

RESULT 14

S45452
inositol-3-phosphate synthase (EC 5.5.1.4) - yeast (*Candida albicans*)
C:Species: *Candida albicans*
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
R:Klig, L.S.; Zobel, P.A.; Devry, C.G.; Losberger, C.
Yeast 10, 789-800, 1994
A:Title: Comparison of INO1 gene sequences and products in *Candida albicans* and *Saccharo*
A:Reference number: S45452; MUID:95065381; PMID:7975896
A:Accession: S45452
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-520 <KLI>
A:Cross-references: UNIPROT:P42800; EMBL:L22737; NID:G413758; P1DN:AAA62849.1; PID:969575;
A:Genetics:
A:Gene: INO1
C:Function: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-
A:Note: requires NAD
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 2.4%; Score 12; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 INGSPTFTVPG 290
|||||
DB 285 INGSPTFTVPG 296

RESULT 15

B48357
Minor outer capsid protein - porcine rotavirus C (strain Cowden)
N:Alternate names: nonstructural protein NS26
C:Species: porcine rotavirus C
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: B48357
R:Bremont, M.; Chabanne-Vautherot, D.; Cohen, J.
Arch. Virol. 130, 85-92, 1993
A:Title: Sequence analysis of three non structural proteins of a porcine group C (Cowden)
A:Reference number: A48357; MUID:93277387; PMID:8389118
A:Accession: B48357
A:Molecule type: genomic RNA
A:Residues: 1-210 <BRES>
A:Cross-references: UNIPROT:P36358
A:Note: sequence extracted from NCBI backbone (NCBIN:132693, NCBI:P132694)
C:Genetics:
A:Map position: segment 10
C:Superfamily: rotavirus minor outer capsid protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:30,120/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.6%; Score 8; DB 1; Length 210;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 313 KMKSVLVD 320
|||||
DB 198 KMKSVLVD 205

Search completed: June 7, 2005, 17:10:36
Job time : 28.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:49:04 ; Search time 86 Seconds
(without alignment)

3036.749 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510

Sequence: 1 MFIENFKVCECPVVKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	510	2	094C02
2	376	73.7	510	2	094K03
3	92	18.0	472	2	094G22
4	75	14.7	510	2	07XJC0
5	70	13.7	510	1	IN01_TOBAC
6	69	13.5	443	2	08H1B7
7	69	13.5	510	1	IN01_NICPA
8	67	13.1	364	2	09AMG8
9	67	13.1	507	1	IN01_CITPA
10	67	13.1	509	2	09AR12
11	67	13.1	510	1	IN01_BRANA
12	67	13.1	510	1	IN01_HORVU
13	67	13.1	510	1	IN01_ORYSA
14	67	13.1	510	1	IN01_SRSIN
15	67	13.1	510	1	IN01_SEIPO
16	67	13.1	510	1	IN01_WHEAT
17	67	13.1	510	1	IN02_ARATH
18	67	13.1	510	1	IN03_ARATH
19	67	13.1	510	2	0944C3
20	67	13.1	510	2	09AV99
21	67	13.1	511	1	IN01_ARATH
22	67	13.1	512	1	IN01_MESCR
23	50	9.8	510	1	IN01_MAIZE
24	48	9.4	511	1	IN01_PHAUV
25	47	9.2	164	2	09FT40
26	43	8.4	512	2	07YZE6
27	38	7.5	186	2	042548
28	36	7.1	409	2	0845A1
29	31	6.1	509	2	0855N2
30	29	5.7	565	1	IN01_DROME
31	27	5.3	135	2	09AT03

32	27	5.3	166	2	09FUP2	09fup2 lycopersico
33	23	4.5	150	2	095U04	095u04 brachiostr
34	23	4.5	560	2	06DDT1	06ddt1 xenopus lae
35	23	4.5	563	2	07ZXU0	07zxy0 xenopus lae
36	21	4.1	561	2	07P2B9	07p2b9 anopheles g
37	17	3.3	526	2	07Z525	07z525 homo sapien
38	17	3.3	557	2	09JHU9	09jhu9 m myo-1nos1
39	17	3.3	558	2	06NXT5	06nxt5 homo sapien
40	17	3.3	558	2	09H2Y2	09h2y2 homo sapien
41	17	3.3	558	2	09NPH2	09nph2 homo sapien
42	17	3.3	558	2	09NHW7	09nwh7 homo sapien
43	16	3.1	220	2	06UP00	06up00 aspergillus
44	16	3.1	537	2	087IU5	087iu5 neurospora
45	15	2.9	181	2	09BT65	09bt65 homo sapien

ALIGNMENTS

RESULT 1	ID	PRELIMINARY;	PRT;	510 AA.
094C02	AC	094C02;		
094C02	DT	01-DEC-2001 (TREMBlrel. 19, Created)		
094C02	DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)		
094C02	DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)		
094C02	DE	Myo-inositol-1-phosphate synthase (EC 5.5.1.4).		
094C02	OS	Glycine max (soybean).		
094C02	OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
094C02	OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
094C02	OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.		
094C02	OX	NCBI_TaxID=3847;		
094C02	RN	[1]		
094C02	RP	SEQUENCE FROM N.A.		
094C02	RA	Hitz W.D., Carlson T.J., Kerr P., Sebastian S.;		
094C02	RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
094C02	RT	EMBL; AY038802; AK72098.1; -.		
094C02	DR	HSSP; P11986; 1P1H.		
094C02	DR	GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.		
094C02	DR	GO; GO:0016853; F:isomerase activity; IEA.		
094C02	DR	GO; GO:0006021; P:myo-inositol biosynthesis; IEA.		
094C02	DR	GO; GO:000654; P:phospholipid biosynthesis; IEA.		
094C02	DR	InterPro: IPR002587; Inos-1-P synth.		
094C02	DR	Pfam: PF01658; Inos-1-P synth; 1.		
094C02	KW	Isomerase.		
094C02	SQ	SEQUENCE 510 AA; 56475 MW; A1B4C77F1643918E CRC64;		
094C02	Query Match	100.0%; Score 510; DB 2; Length 510;		
094C02	Best Local Similarity	100.0%; Pred. No. 0;		
094C02	Matches 510;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
094C02	QY	1 MFIENFKVCECPVVKXTETETI		
094C02	DB	1 MFIENFKVCECPVVKXTETETI		
094C02	QY	61 KLGWLVGVGNGSTLTGCVIANREGISWATKDI		
094C02	DB	61 KLGWLVGVGNGSTLTGCVIANREGISWATKDI		
094C02	QY	121 IYAPFKSLIPVNPDPDIFVGGMDISNMVLADMA		
094C02	DB	121 IYAPFKSLIPVNPDPDIFVGGMDISNMVLADMA		
094C02	QY	181 YDPDFIAANOEERANVIGKTKOEOVOOI		
094C02	DB	181 YDPDFIAANOEERANVIGKTKOEOVOOI		
094C02	QY	241 VGLNTNTENLLAANDREAEISPTLYAIACMENV		
094C02	DB	241 VGLNTNTENLLAANDREAEISPTLYAIACMENV		
094C02	QY	301 LIGGDPFSGOTKMSVVDPLVAGAGIKPTSI		
094C02	DB	301 LIGGDPFSGOTKMSVVDPLVAGAGIKPTSI		

Db 301 LIGGDFKSGQTKMSVLVDPLVGAGIKPTSVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
QY 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKSTIYLAHNTC 420
QY 421 EDSLAAPIIIDVLVLAELSTRIEFKAENEGKFSFHVATILSYLTAKPLVPPTPVN 480
Db 421 EDSLAAPIIIDVLVLAELSTRIEFKAENEGKFSFHVATILSYLTAKPLVPPTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 2

Q94KU3 PRELIMINARY; PRT; 510 AA.
AC 094KU3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Myo-inositol-3-phosphate synthase.
GN Name=MIPS;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
NCBI_TaxID=3847;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21196082; PubMed=11299373; DOI=10.1104/pp.125.4.1941;
RA Hegeman C.E., Good L.L., Grabau B.A.;
RT "Expression of D-myo-inositol-3-phosphate synthase in soybean.
RT Implications for phytic acid biosynthesis.";
RL Plant Physiol. 125:1941-1948 (2001).
DR EMBL; AF293970; AAK49896.1; -.
DR HSSP; P11986; 1PIH.
DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 510 AA; 56506 MW; DE4F3DD7DC6F370 CRC64;

Query Match 73.7%; Score 376; DB 2; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFIENFKVCEPNVKYETEIQSVYNYETTELVEHENGTYOMIVKPKSVKYEFTNIHVP 60
Db 1 MFIENFKVCEPNVKYETEIQSVYNYETTELVEHENGTYOMIVKPKSVKYEFTNIHVP 60
QY 61 KLGVMIVGNGNGSTLTGVIANREGISWATKDKTQQANFGSLTQASAIRVGSFGEE 120
Db 61 KLGVMIVGNGNGSTLTGVIANREGISWATKDKTQQANFGSLTQASAIRVGSFGEE 120
QY 121 IYAPFKSLLPMNPPDIIVFGWDSIMNMLADMAKAVDIDLOKLRPYMESMLPLPGI 180
Db 121 IYAPFKSLLPMNPPDIIVFGWDSIMNMLADMAKAVDIDLOKLRPYMESMLPLPGI 180
QY 181 YPDPFIAAQEERANNVIGTKOEOVOQIIKDKAFKEATKYDKVNVLTANTERSNLY 240
Db 181 YPDPFIAAQEERANNVIGTKOEOVOQIIKDKAFKEATKYDKVNVLTANTERSNLY 240
QY 241 VGLNDTMENILAAVDNEAEISBSTLYAIAQWENVPIFGSPQNTFVGLDLAIARNT 300
Db 241 VGLNDTMENILAAVDNEAEISBSTLYAIAQWENVPIFGSPQNTFVGLDLAIARNT 300
QY 301 LIGGDFKSGQTKMSVLVDPLVGAGIKPTSVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
Db 301 LIGGDFKSGQTKMSVLVDPLVGAGIKPTSVSYNHLGNNDGMNLSAPQTRSKSEIKS 360

QY 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKSTIYLAHNTC 420
QY 421 EDSLAAPIIIDVLVLAELSTRIEFKAENEGKFSFHVATILSYLTAKPLVPPTPVN 477
Db 421 EDSLAAPIIIDVLVLAELSTRIEFKAENEGKFSFHVATILSYLTAKPLVPPTPVN 477

RESULT 3

Q94G22 PRELIMINARY; PRT; 472 AA.
AC 094G22;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE 1L-myo-inositol-1-phosphate synthase (Fragment).
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
NCBI_TaxID=3885;
OX [1]
RN SEQUENCE FROM N.A.
RP Johnson M.D., Lackey K.H., Pope P.M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282263; AAK69514.1; -.
DR HSSP; P11986; 1PIH.
DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
NON TER 472 472
SQ SEQUENCE 472 AA; 52296 MW; 237204E1A370560F CRC64;

Query Match 18.0%; Score 92; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 76-87;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDFKSGQTKMSVLVDPLVGAGIKPTSVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
Db 301 LIGGDFKSGQTKMSVLVDPLVGAGIKPTSVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
QY 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYV 392
Db 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYV 392

RESULT 4

Q7XJC0 PRELIMINARY; PRT; 510 AA.
AC 07XJC0;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Myo-inositol-1-phosphate synthase INOI.
OS Xerophyta viscosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Velloziaceae; Xerophyta.
NCBI_TaxID=90708;
OX [1]
RN SEQUENCE FROM N.A.
RP Majee M., Majumder A.N.L., Mundree S.G.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF323824; AAP85531.1; -.
DR HSSP; P11986; 1JX1.
DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 510 AA; 56234 MW; 5F92212851115A2A CRC64;

Query Match 14.7%; Score 75; DB 2; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4.7e-69;
 Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIYEPHDPHVVKYVPYVDSKRAMDEYTSIFMGKSTIVLHNTCEDSLAAP 428
 DB 369 SNAIYEPHDPHVVKYVPYVDSKRAMDEYTSIFMGKSTIVLHNTCEDSLAAP 428
 QY 429 IILDVLAELSTRI 443
 DB 429 IILDVLAELSTRI 443

RESULT 5

INOL1 TOBAC STANDARD; PRT; 510 AA.
 AC 09LW96;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MT-1-P synthase) (IPS).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 RX MEDLINE=20399434; PubMed=10945337;
 RC TISSUE=Leaf;
 RA Hara K., Yagi M., Koizumi N., Kusano T., Sano H.;
 RT "Screening of wound-responsive genes identifies an immediate-early expressed gene encoding a highly charged protein in mechanically wounded tobacco plants."
 RL Plant Cell Physiol. 41:684-691(2000).
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AB009881; BAA95788.1; -.
 DR HSSP; P11986; 1JXF.
 DR InterPro; IPR002587; Inos-1_P_synth.
 DR Pfam; PF01658; Inos-1_P_synth; 1.
 DR Inositol biosynthesis; isomerase; NAD; phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56369 MW; 4BA6FDDA5DBF6D4D CRC64;

Query Match 13.7%; Score 70; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 8.1e-64;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 RNTLIGDPPKSGQTKMSVLVDFLVGAGIKPTSIYVNLGNNDGMSAPQTFRSKEI 357
 DB 298 RNTLIGDPPKSGQTKMSVLVDFLVGAGIKPTSIYVNLGNNDGMSAPQTFRSKEI 357
 QY 358 SKSNVVDNV 367
 DB 358 SKSNVVDNV 367

RESULT 6

08H1B7 PRELIMINARY; PRT; 443 AA.
 AC 08H1B7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Myo-inositol phosphate synthase (fragment).
 OS Lolium perenne (Perennial ryegrass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae; Pooidae; Poaceae; Lolium.
 NCBI_TaxID=4522;

RN [1]_TaxID=4522;
 RA Amlard V., Prud'homme M.-P., Le Dantec C.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY154382; AAN52772.1; -.
 DR HSSP; P11986; 1JXF.
 DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0006654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 FT NON TER 443
 SQ SEQUENCE 443 AA; 48572 MW; E8BB1A03779B49FA CRC64;

Query Match 13.5%; Score 69; DB 2; Length 443;
 Best Local Similarity 100.0%; Pred. No. 8e-63;
 Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 EGEHDPHVVKYVPYVDSKRAMDEYTSIFMGKSTIVLHNTCEDSLAAPITLDV 434
 DB 375 EGEHDPHVVKYVPYVDSKRAMDEYTSIFMGKSTIVLHNTCEDSLAAPITLDV 434
 QY 435 LLAELSTRI 443
 DB 435 LLAELSTRI 443

RESULT 7

INOL1 NICPA STANDARD; PRT; 510 AA.
 AC 09SSV4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MT-1-P synthase) (IPS).
 GN Name=INPS1.
 OS Nicotiana paniculata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=62141;

RN [1]
 RA Hashimoto A., Yamada S., Komori T.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AB009881; BAA95788.1; -.
 DR HSSP; P11986; 1JXF.
 DR InterPro; IPR002587; Inos-1_P_synth.
 DR Pfam; PF01658; Inos-1_P_synth; 1.
 DR Inositol biosynthesis; isomerase; NAD; phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56369 MW; 4BA6FDDA5DBF6D4D CRC64;

```
CC or send an email to license@sib.ch).
CC -----
DR EMBL: AB032073; BAA84084.1; -.
DR HSSP; P11986; 1PIK.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
KW SEQUENCE 510 AA; 56385 MW; 415B81C27A267666 CRC64;

Query Match
Best Local Similarity 13.5%; Score 69; DB 1; Length 510;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 NTLLGGDDPKSGQTKMSVLDPLVAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 358
DB 299 NTLLGGDDPKSGQTKMSVLDPLVAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 358

QY 359 KSNVVDMMV 367
DB 359 KSNVVDMMV 367

RESULT 8
Q9AMG8 PRELIMINARY; PRT; 364 AA.
AC Q9AMG8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Myo-inositol-1-phosphate synthase.
OS Actinidia arguta.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Ericales; Actinidiaceae; Actinidia.
OX NCBI_TaxID=64478;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Young mature leaf;
RA Klages K., Fitzgerald A., Woodie M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY005128; AAF97409.1; -.
DR HSSP; P11986; 1PIK.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 364 AA; 40246 MW; 1CB64DEFF78127D CRC64;

Query Match
Best Local Similarity 13.1%; Score 67; DB 2; Length 364;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDPKSGQTKMSVLDPLVAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 360
DB 155 LIGGDDPKSGQTKMSVLDPLVAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 214

QY 361 NVVDDMMV 367
DB 215 NVVDDMMV 221

RESULT 9
ID INO1_CITPA STANDARD; PRT; 507 AA.
AC P42802;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
synthase) (Mt-1-P synthase) (IPS).
OC Citrus paradisi (Grapefruit).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=37656;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=leaf;
RA MEDLINE=95148748; PubMed=7846170; DOI=10.1104/pp.106.4.1689;
RX Abu-Abied M., Holland D.;
RT "The gene c-inol from Citrus paradisi is highly homologous to tur1 and
inol from yeast and Spirodela encoding for myo-inositol phosphate
synthase."
RL Plant Physiol. 106:1689-1689 (1994).
CC -1- CATALYTIC ACTIVITY: D-Glucose 6-phosphate = 1D-myo-inositol 3-
phosphate.
CC -1- COPACTOR: NAD.
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL; Z32632; CAA83565.1; -.
DR PIR; S52648; S52648.
DR HSSP; P11986; 1PIK.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 507 AA; 56334 MW; 45D78928991BEDF8 CRC64;

Query Match
Best Local Similarity 13.1%; Score 67; DB 1; Length 507;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDPKSGQTKMSVLDPLVAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 360
DB 301 LIGGDDPKSGQTKMSVLDPLVAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 360

QY 361 NVVDDMMV 367
DB 361 NVVDDMMV 367

RESULT 10
Q9ARI2 PRELIMINARY; PRT; 509 AA.
AC Q9ARI2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Myo-inositol 1-phosphate synthase.
OS Avicennia marina (Grey mangrove).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Acanthaceae; Acanthaceae incertae sedis; Avicennia.
OX NCBI_TaxID=82927;
RN [1]
RP SEQUENCE FROM N.A.
RA Utresh M.N., Parani M., Parida A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028259; AAK21969.1; -.
DR HSSP; P11986; 1PIK.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
```

SQ SEQUENCE 509 AA; 55978 MW; 23C8D354BAF3BD0F CRC64;
 Query Match 13.1%; Score 67; DB 2; Length 509;
 Best Local Similarity 100.0%; Pred. No. 1.1e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIGGDDFKSGQTKMSVLVDPLVAGAGIKPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 300 LIGGDDFKSGQTKMSVLVDPLVAGAGIKPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 359
 QY 361 NVVDDMV 367
 DB 360 NVVDDMV 366

RESULT 11
 INOI BRANA STANDARD; PRT; 510 AA.
 AC 096348;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 OS Brassica napus (Rape).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hussein A., Bourgeois J., Polvi S., Teang E., Keller W.A., Georges F.;
 RT "Cloning of a full length cDNA encoding myo-inositol 1-phosphate synthase from Brassica napus."
 RT Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U66307; AAB06756.2; -.
 DR PIR; T08436; T08436.
 DR HSSP; P11986; P119.
 DR InterPro; IPR002587; Inos-1-P synth.
 DR Pfam; PF01658; Inos-1-P synth. 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56377 MW; A40EB6558D80739 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.1e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGAGIKPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 301 LIGGDDFKSGQTKMSVLVDPLVAGAGIKPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 12

INOI HORVU STANDARD; PRT; 510 AA.
 AC 065195;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, Harrington;
 RA Larson S.R., Raboy V.;
 RT "Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes."
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF056325; AAC17133.1; -.
 DR PIR; T04399; T04399.
 DR HSSP; P11986; P119.
 DR InterPro; IPR002587; Inos-1-P synth.
 DR Pfam; PF01658; Inos-1-P synth. 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56173 MW; EA63138121692724 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.1e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGAGIKPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 301 LIGGDDFKSGQTKMSVLVDPLVAGAGIKPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 13
 INOI ORYSA STANDARD; PRT; 510 AA.
 AC 064437;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 GN Name=INOI;
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;


```

RP SEQUENCE FROM N.A.
RC STRAIN=cnv. Japonica / Kamenoo;
RA Yoshida K.T., Wada T., Koyama H., Mizobuchi-Fukunaka R., Naito S.;
RT "Expression of myo-inositol 1-phosphate synthase gene and phytin
RL accumulation during seed development in rice.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB012107; BAA25729.1; -.
DR HSSP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth.1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56216 MW; 2580220DD871A80 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.1e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDFKSGGQTKMSVLVDPLVGAGIKPTISIVSYNHLGNNDGNNLSAPQTFRSKEISK 360
DB 301 LIGGDFKSGGQTKMSVLVDPLVGAGIKPTISIVSYNHLGNNDGNNLSAPQTFRSKEISK 360
QY 361 NVVDDMV 367
DB 361 NVVDDMV 367

RESULT 14
INOL SESIN STANDARD; PRT; 510 AA.
AC Q9RYV1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (MI-1-P synthase) (IPS).
OS Sesamum indicum (Oriental sesame) (gingelly).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiales; Pedaliaceae; Sesamum.
OC NCBI_Taxid=4182;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RA Jin U.-H., Chung C.-H.;
RT "Characterization and functional analysis of a myo-inositol 1-
RT phosphate synthase cDNA from sesame (Sesamum indicum L.) seeds.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF284065; AAC01148.1; -.
DR HSSP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth.1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56214 MW; 88D7376CEB73401F CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.1e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDFKSGGQTKMSVLVDPLVGAGIKPTISIVSYNHLGNNDGNNLSAPQTFRSKEISK 360
DB 301 LIGGDFKSGGQTKMSVLVDPLVGAGIKPTISIVSYNHLGNNDGNNLSAPQTFRSKEISK 360
QY 361 NVVDDMV 367
DB 361 NVVDDMV 367

RESULT 15
INOL SPIPO STANDARD; PRT; 510 AA.
AC P42803;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (MI-1-P synthase) (IPS).
GN Name=SPIR1;
OS Spirodela polyrrhiza (giant duckweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Lemnoidae;
OC Spirodela.
OC NCBI_Taxid=29656;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94035182; PubMed=8220483;
RA Smart C.C., Fleming A.U.;
RT "A plant gene with homology to D-myo-inositol-3-phosphate synthase is
RT rapidly and spatially up-regulated during an abscisic-acid-induced
RL morphogenic response in Spirodela polyrrhiza.";
RL Plant J. 4:279-293(1993).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- INDUCTION: By abscisic acid (ABA).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11693; CAA7751.1; -.
DR PIR; S60302; S60302.
DR HSSP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth.1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.

```

SQ SEQUENCE 510 AA; 56385 MW; 2D56D366FC5E03C CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;

Best Local Similarity 100.0%; Pred.No. 1.1e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDPKSGQTMKSVLVDFLVGAGIKPTSIYSYNNHNGNDGMLSAFOTFRSKETSKS 360

DB 301 LIGGDDPKSGQTMKSVLVDFLVGAGIKPTSIYSYNNHNGNDGMLSAFOTFRSKETSKS 360

QY 361 NVVDDMV 367

DB 361 NVVDDMV 367

Search completed: June 7, 2005, 17:09:33

Job time : 88 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:00:20 ; Search time 31 Seconds
(without alignments)
1228.098 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MFIENPKVCEPDKVKTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgnt2_6/prodata/1/iaa/5A_COMB.pep:*
- 2: /cgnt2_6/prodata/1/iaa/5B_COMB.pep:*
- 3: /cgnt2_6/prodata/1/iaa/6A_COMB.pep:*
- 4: /cgnt2_6/prodata/1/iaa/6B_COMB.pep:*
- 5: /cgnt2_6/prodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgnt2_6/prodata/1/iaa/backfile01.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	13.3	510	4	US-09-727-628-2
2	67	13.1	510	3	US-09-118-442-11
3	67	13.1	510	3	US-09-677-064-11
4	13	2.5	533	4	US-09-734-237B-73
5	13	2.5	534	4	US-09-734-237B-75
6	12	2.4	525	4	US-09-248-796A-17234
7	9	1.8	1481	4	US-09-231-899-70
8	7	1.4	55	4	US-09-640-211A-2116
9	7	1.4	79	4	US-09-107-433-4073
10	7	1.4	93	4	US-09-502-540-15484
11	7	1.4	134	4	US-09-270-767-34609
12	7	1.4	134	4	US-09-270-767-49826
13	7	1.4	160	4	US-09-957-641A-19
14	7	1.4	191	4	US-09-252-991A-24024
15	7	1.4	224	4	US-08-372-255-16
16	7	1.4	224	5	PCT-US95-08565-16
17	7	1.4	254	4	US-09-252-991A-21339
18	7	1.4	269	4	US-09-902-540-16003
19	7	1.4	273	3	US-08-235-836C-142
20	7	1.4	273	3	US-08-235-836C-144
21	7	1.4	284	4	US-09-914-259-62
22	7	1.4	284	4	US-10-164-595-32
23	7	1.4	295	4	US-09-270-767-57050
24	7	1.4	295	4	US-09-270-767-57401
25	7	1.4	306	4	US-09-107-532A-5995
26	7	1.4	308	4	US-09-489-039A-11598
27	7	1.4	322	4	US-09-565-501A-112

28	7	1.4	322	4	US-09-639-206A-112	Sequence 112, App
29	7	1.4	322	4	US-09-874-923-112	Sequence 112, App
30	7	1.4	328	4	US-09-583-110-4209	Sequence 4209, Ap
31	7	1.4	340	4	US-09-489-039A-9804	Sequence 9804, Ap
32	7	1.4	360	4	US-09-489-039A-9936	Sequence 9936, Ap
33	7	1.4	361	4	US-09-710-279-160	Sequence 160, App
34	7	1.4	364	4	US-09-902-540-16435	Sequence 16435, A
35	7	1.4	381	4	US-09-254-776B-53	Sequence 53, Appl
36	7	1.4	388	4	US-09-949-016-7631	Sequence 7631, Ap
37	7	1.4	399	4	US-09-543-681A-5325	Sequence 5325, Ap
38	7	1.4	443	4	US-09-270-767-46616	Sequence 46616, A
39	7	1.4	443	4	US-09-949-016-11613	Sequence 11613, A
40	7	1.4	446	4	US-09-252-991A-22844	Sequence 22844, A
41	7	1.4	456	4	US-09-328-352-6946	Sequence 6946, Ap
42	7	1.4	467	4	US-09-489-039A-12726	Sequence 12726, A
43	7	1.4	472	4	US-09-489-039A-8121	Sequence 8121, Ap
44	7	1.4	472	4	US-09-134-000C-3893	Sequence 3893, Ap
45	7	1.4	548	4	US-09-902-540-11870	Sequence 11870, A

ALIGNMENTS

```

RESULT 1
US-09-727-628-2
; Sequence 2, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727,628
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/166,612
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-727-628-2

Query Match      13.3%; Score 68; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.1e-62;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      376 PSHPHVVVIVKVPVYVGSKRAMEYTSIFMGKSTVLNHTCEDSLAAPPITLDVL 435
      |||
Db      376 PSHPHVVVIVKVPVYVGSKRAMEYTSIFMGKSTVLNHTCEDSLAAPPITLDVL 435
      |||

OY      436 LABELSTRI 443
      |||
Db      436 LABELSTRI 443
      |||

RESULT 2
US-09-118-442-11
; Sequence 11, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catv, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; FILE REFERENCE: 0706

```

```
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ CURRENT FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 510
/ TYPE: PRF
/ ORGANISM: Zea mays
US-09-118-442-11
```

```
Query Match      13.1%; Score 67; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.4e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 301 LGGDDFKSGQTKMSVLVDPLVAGIKPTISVYNHLNNDGMNLSAPQTFRSKXISKS 360
DB 301 LGGDDFKSGQTKMSVLVDPLVAGIKPTISVYNHLNNDGMNLSAPQTFRSKXISKS 360
```

```
QY 361 NVVDDMV 367
DB 361 NVVDDMV 367
```

```
RESULT 3
US-09-677-064-11
/ Sequence 11, Application US/09677064
/ Patent No. 6291224
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677,064
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ PRIOR FILING DATE: 1998-07-17
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 510
/ TYPE: PRF
/ ORGANISM: Zea mays
US-09-677-064-11
```

```
Query Match      13.1%; Score 67; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.4e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 301 LGGDDFKSGQTKMSVLVDPLVAGIKPTISVYNHLNNDGMNLSAPQTFRSKXISKS 360
DB 301 LGGDDFKSGQTKMSVLVDPLVAGIKPTISVYNHLNNDGMNLSAPQTFRSKXISKS 360
```

```
QY 361 NVVDDMV 367
DB 361 NVVDDMV 367
```

```
RESULT 4
US-09-734-237B-73
```

```
/ Sequence 73, Application US/09734237B
/ Patent No. 6818752
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bul, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ CURRENT FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 73
/ LENGTH: 533
/ TYPE: PRF
/ ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-73
```

```
Query Match      2.5%; Score 13; DB 4; Length 533;
Best Local Similarity 100.0%; Pred. No. 9.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 279 INGSPONTFVPG 291
DB 293 INGSPONTFVPG 305
```

```
RESULT 5
US-09-734-237B-75
/ Sequence 75, Application US/09734237B
/ Patent No. 6818752
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bul, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ CURRENT FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 75
/ LENGTH: 534
/ TYPE: PRF
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic protein derived from Saccharomyces cerevisiae myo-inosi
/ OTHER INFORMATION: tol-1-phosphate synthase, having a glycine residue inserted after
US-09-734-237B-75
```

```
Query Match      2.5%; Score 13; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 9.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 279 INGSPONTFVPG 291
DB 294 INGSPONTFVPG 306
```

```
RESULT 6
US-09-248-796A-17234
/ Sequence 17234, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstock et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
/ FILE REFERENCE: 107196.132
```

;; CURRENT APPLICATION NUMBER: US/09/248,796A
;; CURRENT FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074,725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096,409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 17234
;; LENGTH: 525
;; TYPE: PRT
;; ORGANISM: Candida albicans
US-09-248-796A-17234

Query Match 2.4%; Score 12; DB 4; Length 525;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 INGSPTFTVPG 290
Db 290 INGSPTFTVPG 301

RESULT 7
US-09-231-899-70
; Sequence 70, Application US/09231899
; Patent No. 656583
; GENERAL INFORMATION:
; APPLICANT: Lasmer, Michael
; APPLICANT: Metz, James G
; APPLICANT: Facciotti, Daniel
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKS GENES
; CURRENT APPLICATION NUMBER: US/09/231,899
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: 09/090,793
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 1481
; TYPE: PRT
; ORGANISM: Schizochytrium aggregatum
US-09-231-899-70

Query Match 1.8%; Score 9; DB 4; Length 1481;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 ENLLAAVDR 256
Db 1092 ENLLAAVDR 1100

RESULT 8
US-09-640-211A-2116
; Sequence 2116, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2116
; LENGTH: 55

;; TYPE: PRT
;; ORGANISM: Eucalyptus grandis
US-09-640-211A-2116

Query Match 1.4%; Score 7; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 485 ORAMLEN 491
Db 36 ORAMLEN 42

RESULT 9
US-09-107-433-4073
; Sequence 4073, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4073:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...79
; SEQUENCE DESCRIPTION: SEQ ID NO: 4073:
US-09-107-433-4073

Query Match 1.4%; Score 7; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 444 EFKAEKE 450
Db 71 EFKAEKE 77

```
RESULT 10
US-09-902-540-15484
; Sequence 15484, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15484
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15484

Query Match      1.4%; Score 7; DB 4; Length 93;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      473 TPVPGI 479      |||||
DB      54 TPVPGI 60

RESULT 11
US-09-270-767-34609
; Sequence 34609, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34609
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34609

Query Match      1.4%; Score 7; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      286 TPVPGI 292      |||||
DB      99 TPVPGI 105

RESULT 12
US-09-270-767-49826
; Sequence 49826, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
```

```
; SEQ ID NO 49826
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49826

Query Match      1.4%; Score 7; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      286 TPVPGI 292      |||||
DB      99 TPVPGI 105

RESULT 13
US-09-957-641A-19
; Sequence 19, Application US/09957641A
; Patent No. 6770744
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-00 US
; CURRENT APPLICATION NUMBER: US/09/957,641A
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/234,047
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/236,460
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Porcine
US-09-957-641A-19

Query Match      1.4%; Score 7; DB 4; Length 160;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      476 TPVNAL 482      |||||
DB      119 TPVNAL 125

RESULT 14
US-09-252-991A-24024
; Sequence 24024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24024
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24024

Query Match      1.4%; Score 7; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 469 APLVPPG 475
 Db 85 APLVPPG 91

RESULT 15

US-08-272-255-16
 ; Sequence 16, Application US/08272255
 ; Patent No. 5824859
 ; GENERAL INFORMATION:
 ; APPLICANT: Cashmore, Anthony R.
 ; APPLICANT: Ahmad, Margaret
 ; APPLICANT: Lin, Chentao
 ; TITLE OF INVENTION: Blue light Photoreceptors and Methods of
 ; TITLE OF INVENTION: Using the same
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859r1s
 ; STREET: One Liberty Place, 46th floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/272,255
 ; FILING DATE: 08-JUL-1994
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Leary Ph.D., Kathryn
 ; REGISTRATION NUMBER: 36,317
 ; REFERENCE/DOCKET NUMBER: UPN-1795
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-3100
 ; TELEFAX: (215) 568-3439
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 224 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-272-255-16

Query Match 1.4%; Score 7; DB 2; Length 224;
 Best Local Similarity 100.0%; Pred.No. 81;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 421 EDSLILAA 427
 Db 8 EDSLILAA 14

Search completed: June 7, 2005, 17:11:47
 Job time : 33 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:09:41, Search time 77.5 Seconds
(without alignments)
2370.369 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510

Sequence: 1 MFIENPKVCECPVVKYETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 1599520 seqs, 360203123 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database: Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubppa/US11A_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	510	14	US-10-025-003-2
2	510	100.0	510	14	US-10-025-003-10
3	510	100.0	510	16	US-10-718-952-2
4	510	100.0	510	16	US-10-718-952-10
5	440	86.3	511	15	US-10-424-599-213009
6	409	80.2	510	14	US-10-025-003-6
7	409	80.2	510	16	US-10-718-952-6
8	253	49.6	510	14	US-10-025-003-14
9	253	49.6	510	14	US-10-025-003-16
10	253	49.6	510	16	US-10-718-952-14
11	253	49.6	510	16	US-10-718-952-16
12	236	46.3	510	14	US-10-025-003-12

13	236	46.3	510	16	US-10-718-952-12	Sequence 12, Appl
14	159	31.2	431	15	US-10-424-599-154863	Sequence 154863,
15	113	22.2	220	15	US-10-424-599-259439	Sequence 259439,
16	103	20.2	211	15	US-10-424-599-165505	Sequence 165505,
17	102	20.0	510	15	US-10-424-599-154864	Sequence 154864,
18	73	14.3	250	15	US-10-424-599-154861	Sequence 154861,
19	68	13.3	332	15	US-10-425-114-42702	Sequence 42702, A
20	68	13.3	394	16	US-10-363-829-458	Sequence 458, App
21	68	13.3	450	16	US-10-425-115-231845	Sequence 231845,
22	68	13.3	510	9	US-09-727-628-2	Sequence 2, Appl1
23	68	13.3	510	16	US-10-425-115-231850	Sequence 231850,
24	68	13.3	510	16	US-10-425-115-231853	Sequence 231853,
25	68	13.3	510	16	US-10-425-115-231857	Sequence 231857,
26	68	13.3	512	15	US-10-425-114-39621	Sequence 39621, A
27	68	13.3	512	15	US-10-425-114-46916	Sequence 46916, A
28	68	13.3	512	15	US-10-425-114-53231	Sequence 53231, A
29	68	13.3	512	15	US-10-425-114-66216	Sequence 66216, A
30	67	13.1	458	16	US-10-425-115-231843	Sequence 231843,
31	67	13.1	505	15	US-10-442-017-15	Sequence 15, Appl
32	67	13.1	510	9	US-09-921-232-11	Sequence 11, Appl
33	67	13.1	510	9	US-09-921-330-11	Sequence 11, Appl
34	67	13.1	510	9	US-09-921-329-11	Sequence 11, Appl
35	67	13.1	510	16	US-10-437-963-173252	Sequence 173252,
36	67	13.1	510	16	US-10-767-701-46278	Sequence 46278,
37	67	13.1	510	16	US-10-425-115-231852	Sequence 231852,
38	67	13.1	510	16	US-10-425-115-306814	Sequence 306814,
39	67	13.1	510	16	US-10-739-930-7635	Sequence 7635, Ap
40	67	13.1	516	15	US-10-425-114-62568	Sequence 62568, A
41	67	13.1	542	15	US-10-425-114-58674	Sequence 58674, A
42	66	12.9	124	15	US-10-424-599-213004	Sequence 213004,
43	47	9.2	78	15	US-10-424-599-213008	Sequence 213008,
44	42	8.2	84	15	US-10-424-599-154859	Sequence 154859,
45	31	6.1	100	16	US-10-425-115-231841	Sequence 231841,

ALIGNMENTS

RESULT 1
US-10-025-003-2
; Sequence 2, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strell, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-2

Query Match 100.0%; Score 510; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFIENPKVCECPVVKYETETIETELVHNRNGYQWIVVVKYSKYETETNIHVP 60
DB 1 MFIENPKVCECPVVKYETETIETELVHNRNGYQWIVVVKYSKYETETNIHVP 60
QY 61 KLGWMLVGMGNGNGSTLTCGVANRNGISWATKDKIQQANVYFGSLQSAIVGSGQBE 120

```
Db 61 KLGWMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Qy 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKAVFDILOQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKAVFDILOQLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANOEBRANNAVIGTKQEOVOQI IKDKAFKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPDFIAANOEBRANNAVIGTKQEOVOQI IKDKAFKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPIINGSPONTFVPGILDLAIARNT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPIINGSPONTFVPGILDLAIARNT 300
Qy 301 LIGGDDFKSGGOTKMSVVDPLVGAGIKPTSIIVSNH LGNDGMNLSAPOTFRSEISKS 360
Db 301 LIGGDDFKSGGOTKMSVVDPLVGAGIKPTSIIVSNH LGNDGMNLSAPOTFRSEISKS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVVKVPPYVGD SKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVVKVPPYVGD SKRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIEFKAENEGKFSFHVATILSYLTAKAPLVPPTPVVN 480
Db 421 EDSLAAPIIIDLVLLAELSTRIEFKAENEGKFSFHVATILSYLTAKAPLVPPTPVVN 480
Qy 481 ALSKORAMLENTMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTMRACVGLAPENNMILEYK 510
```

RESULT 2

```
US-10-025-003-10
; Sequence 10, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRF
; ORGANISM: Glycine max
US-10-025-003-10
```

```
Query Match 100.0%; Score 510; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFIENFKVCEPNVKYETEIGSVYNYETTELVEHNRNGTYOWIVPKSVKTEFKTNIHVP 60
Db 1 MFIENFKVCEPNVKYETEIGSVYNYETTELVEHNRNGTYOWIVPKSVKTEFKTNIHVP 60
Qy 61 KLGWMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Db 61 KLGWMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Qy 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKAVFDILOQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKAVFDILOQLRPYMESMLPLPGI 180
```

```
Qy 181 YDPDFIAANOEBRANNAVIGTKQEOVOQI IKDKAFKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPDFIAANOEBRANNAVIGTKQEOVOQI IKDKAFKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPIINGSPONTFVPGILDLAIARNT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPIINGSPONTFVPGILDLAIARNT 300
Qy 301 LIGGDDFKSGGOTKMSVVDPLVGAGIKPTSIIVSNH LGNDGMNLSAPOTFRSEISKS 360
Db 301 LIGGDDFKSGGOTKMSVVDPLVGAGIKPTSIIVSNH LGNDGMNLSAPOTFRSEISKS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVVKVPPYVGD SKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVVKVPPYVGD SKRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIEFKAENEGKFSFHVATILSYLTAKAPLVPPTPVVN 480
Db 421 EDSLAAPIIIDLVLLAELSTRIEFKAENEGKFSFHVATILSYLTAKAPLVPPTPVVN 480
Qy 481 ALSKORAMLENTMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTMRACVGLAPENNMILEYK 510
```

RESULT 3

```
US-10-718-952-2
; Sequence 2, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRF
; ORGANISM: Glycine max
US-10-718-952-2
```

```
Query Match 100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFIENFKVCEPNVKYETEIGSVYNYETTELVEHNRNGTYOWIVPKSVKTEFKTNIHVP 60
Db 1 MFIENFKVCEPNVKYETEIGSVYNYETTELVEHNRNGTYOWIVPKSVKTEFKTNIHVP 60
Qy 61 KLGWMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Db 61 KLGWMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Qy 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKAVFDILOQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKAVFDILOQLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANOEBRANNAVIGTKQEOVOQI IKDKAFKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPDFIAANOEBRANNAVIGTKQEOVOQI IKDKAFKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPIINGSPONTFVPGILDLAIARNT 300
```

```
Db 241 VGNLNTMENLLAAYDRNEAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNT 300
Qy 301 LIGGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNNDGMNLAPQTFRSKEISKS 360
Db 301 LIGGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNNDGMNLAPQTFRSKEISKS 360
Qy 361 NVVDDMVNSNALIYEGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLAEIESTRIEFAKNEGKFHSPVATILSYLTKAPLVPGTPPVN 480
Db 421 EDSLAAPIIIDLVLAEIESTRIEFAKNEGKFHSPVATILSYLTKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 4

```
US-10-718-952-10
; Sequence 10, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-10
```

```
Query Match 100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MEIENFKVECPNVKTYETETIOSVYNYETTELVEHNNGTYYQWIVKPKSVKYEFTNIHVP 60
Db 1 MEIENFKVECPNVKTYETETIOSVYNYETTELVEHNNGTYYQWIVKPKSVKYEFTNIHVP 60
Qy 61 KGVNMLVKGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGREE 120
Db 61 KGVNMLVKGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGREE 120
Qy 121 IYAPFESKLLPMVNPDDIVEFGWDISNMNLADAMARAKVFDIDLQKQLRPPYMSMLPLPPI 180
Db 121 IYAPFESKLLPMVNPDDIVEFGWDISNMNLADAMARAKVFDIDLQKQLRPPYMSMLPLPPI 180
Qy 181 YPPDFLANQEBRANNVYGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERYSNLV 240
Db 181 YPPDFLANQEBRANNVYGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERYSNLV 240
Qy 241 VGNLNTMENLLAAYDRNEAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNT 300
Db 241 VGNLNTMENLLAAYDRNEAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNT 300
Qy 301 LIGGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNNDGMNLAPQTFRSKEISKS 360
Db 301 LIGGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNNDGMNLAPQTFRSKEISKS 360
```

```
Qy 361 NVVDDMVNSNALIYEGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLAEIESTRIEFAKNEGKFHSPVATILSYLTKAPLVPGTPPVN 480
Db 421 EDSLAAPIIIDLVLAEIESTRIEFAKNEGKFHSPVATILSYLTKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 5

```
US-10-424-599-213009
; Sequence 213009, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 213009
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(511)
; OTHER INFORMATION: unsure at all Xaa Locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.pep
US-10-424-599-213009
```

```
Query Match 86.3%; Score 440; DB 15; Length 511;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 71 GNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGEEIYAPFESKLLP 130
Db 71 GNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGEEIYAPFESKLLP 130
Qy 131 MNPPDDIVEFGWDISNMNLADAMARAKVFDIDLQKQLRPPYMSMLPLPPIYPPDFLANQ 190
Db 131 MNPPDDIVEFGWDISNMNLADAMARAKVFDIDLQKQLRPPYMSMLPLPPIYPPDFLANQ 190
Qy 191 EERANNVYGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERYSNLVVGLNTMENTL 250
Db 191 EERANNVYGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERYSNLVVGLNTMENTL 250
Qy 251 LAAVDRNEAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNTLIGGDDPKSG 310
Db 251 LAAVDRNEAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNTLIGGDDPKSG 310
Qy 311 QTKMSVLDPLVGAGIKPTSIYSYNHLGNNDGMNLAPQTFRSKEISKNVDDMVNSN 370
Db 311 QTKMSVLDPLVGAGIKPTSIYSYNHLGNNDGMNLAPQTFRSKEISKNVDDMVNSN 370
Qy 371 ALIYEPGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTCEDSLAAPII 430
Db 371 ALIYEPGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTCEDSLAAPII 430
Qy 431 IDLVLLAEIESTRIEFAKNEGKFHSPVATILSYLTKAPLVPGTPPVNALSOKRAMLE 490
Db 431 IDLVLLAEIESTRIEFAKNEGKFHSPVATILSYLTKAPLVPGTPPVNALSOKRAMLE 490
```

QY 491 NIMRACVGLAPENNMILEYK 510
Db 492 NIMRACVGLAPENNMILEYK 511

RESULT 6
US-10-025-003-6
; Sequence 6, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-6

Query Match 80.2%; Score 409; DB 14; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MFENFKVCEPNVKYTEIEIOSVYNYETTELVEHNRNTGYMIVKPKSVKTEFKTNHVP 60
Db 1 MFENFKVCEPNVKYTEIEIOSVYNYETTELVEHNRNTGYMIVKPKSVKTEFKTNHVP 60
QY 61 KLGWLVGNGNGSTLTGVIANREGISWATKDKIQOANFGSLTQASAIRVGSFGQEE 120
Db 61 KLGWLVGNGNGSTLTGVIANREGISWATKDKIQOANFGSLTQASAIRVGSFGQEE 120
QY 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YPDPFIANOEERANNVIGTKOEVOQI IKDIAFKKATKYDKVVTMTANTERYSNLV 240
Db 181 YPDPFIANOEERANNVIGTKOEVOQI IKDIAFKKATKYDKVVTMTANTERYSNLV 240
QY 241 VGLNDTMENLLAANDRNEABISPTLYAIACVMEVPIFGSPONTFVPGGLIDLAIAANT 300
Db 241 VGLNDTMENLLAANDRNEABISPTLYAIACVMEVPIFGSPONTFVPGGLIDLAIAANT 300
QY 301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRANDEYTSIEIFMGKSTIYLANTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRANDEYTSIEIFMGKSTIYLANTC 420
QY 421 EDSLLAAPILIDVLVLAELSTRIEFKANEGKFSFHVATILSYLTAKPLVPGTPVYN 480
Db 421 EDSLLAAPILIDVLVLAELSTRIEFKANEGKFSFHVATILSYLTAKPLVPGTPVYN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 7

US-10-718-952-6
; Sequence 6, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-6

Query Match 80.2%; Score 409; DB 16; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MFENFKVCEPNVKYTEIEIOSVYNYETTELVEHNRNTGYMIVKPKSVKTEFKTNHVP 60
Db 1 MFENFKVCEPNVKYTEIEIOSVYNYETTELVEHNRNTGYMIVKPKSVKTEFKTNHVP 60
QY 61 KLGWLVGNGNGSTLTGVIANREGISWATKDKIQOANFGSLTQASAIRVGSFGQEE 120
Db 61 KLGWLVGNGNGSTLTGVIANREGISWATKDKIQOANFGSLTQASAIRVGSFGQEE 120
QY 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YPDPFIANOEERANNVIGTKOEVOQI IKDIAFKKATKYDKVVTMTANTERYSNLV 240
Db 181 YPDPFIANOEERANNVIGTKOEVOQI IKDIAFKKATKYDKVVTMTANTERYSNLV 240
QY 241 VGLNDTMENLLAANDRNEABISPTLYAIACVMEVPIFGSPONTFVPGGLIDLAIAANT 300
Db 241 VGLNDTMENLLAANDRNEABISPTLYAIACVMEVPIFGSPONTFVPGGLIDLAIAANT 300
QY 301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRANDEYTSIEIFMGKSTIYLANTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRANDEYTSIEIFMGKSTIYLANTC 420
QY 421 EDSLLAAPILIDVLVLAELSTRIEFKANEGKFSFHVATILSYLTAKPLVPGTPVYN 480
Db 421 EDSLLAAPILIDVLVLAELSTRIEFKANEGKFSFHVATILSYLTAKPLVPGTPVYN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 8
US-10-025-003-14
; Sequence 14, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott

APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 14
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-14

Query Match 49.6%; Score 253; DB 14; Length 510;
Best Local Similarity 99.7%; Pred. No. 7.9e-242;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HYPKGLVNLVWGNGNSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
DB 58 HYPKGLVNLVWGNGNSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
QY 118 GEEIYAPFKSLIPWNPDDIVFGGWDISMNMLADAMARAKVPDIDLOKOLRPYMSMLPL 177
DB 118 GEEIYAPFKSLIPWNPDDIVFGGWDISMNMLADAMARAKVPDIDLOKOLRPYMSMLPL 177
QY 178 PGTYDPDFIAANOEBERANNVIKGTQEOVOQIHKIKAFKATKVDKVVLTANTERS 237
DB 178 PGTYDPDFIAANOEBERANNVIKGTQEOVOQIHKIKAFKATKVDKVVLTANTERS 237
QY 238 NLVGLNDTMENTLLAADVDRNEAISPSTLYALACWENVPFINGSQNTFVGLIDLAL 297
DB 238 NLVGLNDTMENTLLAADVDRNEAISPSTLYALACWENVPFINGSQNTFVGLIDLAL 297
QY 298 RNTLIGDDFKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
DB 298 RNTLIGDDFKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
QY 358 SSKSNVDDMVNSNALIYEPGEHPDHVVYIKVPYVGDSCRANDEYTSIFMGK 411
DB 358 SSKSNVDDMVNSNALIYEPGEHPDHVVYIKVPYVGDSCRANDEYTSIFMGK 411

RESULT 9

US-10-025-003-16
Sequence 16, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 16
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-16

Query Match 49.6%; Score 253; DB 14; Length 510;
Best Local Similarity 99.7%; Pred. No. 7.9e-242;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HYPKGLVNLVWGNGNSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
DB 58 HYPKGLVNLVWGNGNSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
QY 118 GEEIYAPFKSLIPWNPDDIVFGGWDISMNMLADAMARAKVPDIDLOKOLRPYMSMLPL 177
DB 118 GEEIYAPFKSLIPWNPDDIVFGGWDISMNMLADAMARAKVPDIDLOKOLRPYMSMLPL 177
QY 178 PGTYDPDFIAANOEBERANNVIKGTQEOVOQIHKIKAFKATKVDKVVLTANTERS 237
DB 178 PGTYDPDFIAANOEBERANNVIKGTQEOVOQIHKIKAFKATKVDKVVLTANTERS 237
QY 238 NLVGLNDTMENTLLAADVDRNEAISPSTLYALACWENVPFINGSQNTFVGLIDLAL 297
DB 238 NLVGLNDTMENTLLAADVDRNEAISPSTLYALACWENVPFINGSQNTFVGLIDLAL 297
QY 298 RNTLIGDDFKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
DB 298 RNTLIGDDFKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
QY 358 SSKSNVDDMVNSNALIYEPGEHPDHVVYIKVPYVGDSCRANDEYTSIFMGK 411
DB 358 SSKSNVDDMVNSNALIYEPGEHPDHVVYIKVPYVGDSCRANDEYTSIFMGK 411

RESULT 10

US-10-718-952-14
Sequence 14, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 14
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-718-952-14

Query Match 49.6%; Score 253; DB 16; Length 510;
Best Local Similarity 99.7%; Pred. No. 7.9e-242;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HYPKGLVNLVWGNGNSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
DB 58 HYPKGLVNLVWGNGNSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
QY 118 GEEIYAPFKSLIPWNPDDIVFGGWDISMNMLADAMARAKVPDIDLOKOLRPYMSMLPL 177
DB 118 GEEIYAPFKSLIPWNPDDIVFGGWDISMNMLADAMARAKVPDIDLOKOLRPYMSMLPL 177
QY 178 PGTYDPDFIAANOEBERANNVIKGTQEOVOQIHKIKAFKATKVDKVVLTANTERS 237
DB 178 PGTYDPDFIAANOEBERANNVIKGTQEOVOQIHKIKAFKATKVDKVVLTANTERS 237
QY 238 NLVGLNDTMENTLLAADVDRNEAISPSTLYALACWENVPFINGSQNTFVGLIDLAL 297


```
|||||
Db 238 NLVGLNDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 297
Qy 298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
Db 298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
Qy 358 SKSNVVDMDVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
Db 358 SKSNVVDMDVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
```

```
RESULT 11
US-10-718-952-16
; Sequence 16, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-16
```

```
Query Match 49.6%; Score 253; DB 16; Length 510;
Best Local Similarity 99.7%; Pred. No. 7.9e-242;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 58 HVPKLGWMLVGVGNNGSTLTGGVIANREGISMATKDKIQANYPFGLTQASAIRVGSFO 117
Db 58 HVPKLGWMLVGVGNNGSTLTGGVIANREGISMATKDKIQANYPFGLTQASAIRVGSFO 117
Qy 118 GEETIAPFSSLPMTVPDDIVFGMDISNMNLADAAKAKVPDIDQKOLRPFMESMPL 177
Db 118 GEETIAPFSSLPMTVPDDIVFGMDISNMNLADAAKAKVPDIDQKOLRPFMESMPL 177
Qy 178 PGTYDDPFLAANOEBEERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTERS 237
Db 178 PGTYDDPFLAANOEBEERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTERS 237
Qy 238 NLVGLNDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 297
Db 238 NLVGLNDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 297
Qy 298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
Db 298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
Qy 358 SKSNVVDMDVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
Db 358 SKSNVVDMDVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
```

```
RESULT 12
US-10-025-003-12
; Sequence 12, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
```

```
/ APPLICANT: Sebastian, Scott
/ APPLICANT: Grace, John
/ APPLICANT: Streitz, Leon
/ TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
/ TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
/ FILE REFERENCE: BB-1077-C
/ CURRENT APPLICATION NUMBER: US/10/025,003
/ CURRENT FILING DATE: 2002-05-07
/ PRIOR APPLICATION NUMBER: 08/835,751
/ PRIOR FILING DATE: APRIL 8, 1997
/ PRIOR APPLICATION NUMBER: PCT/US98/06822
/ PRIOR FILING DATE: APRIL 7, 1998
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 12
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Glycine max
US-10-025-003-12
```

```
Query Match 46.3%; Score 236; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 5.8e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 PLPGTYDDPFLAANOEBEERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTER 235
Db 176 PLPGTYDDPFLAANOEBEERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTER 235
Qy 236 YSNLVGNDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 295
Db 236 YSNLVGNDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 295
Qy 296 IARNTLIGDDPKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSK 355
Db 296 IARNTLIGDDPKSGQTKMSVLVDPLVAGIKRPTSIYSYNHLGNNDGMNLSAPQTFRSK 355
Qy 356 EISKSNNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
Db 356 EISKSNNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
```

```
RESULT 13
US-10-718-952-12
; Sequence 12, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-12
```

```
Query Match 46.3%; Score 236; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 5.8e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 PLPGTYDDPFLAANOEBEERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTER 235
```

Db 176 PLPGIYDPEPIAANGERRANNVIKGRKQEQVQIIKDIKAPKATKVDKVVLLMTANTER 235

Qy 236 YSLNVLVGLNDTMENLLAAVDRNEAEISPSITLAIACWENVPFINGSPQNTFVGLIDLA 295

Db 236 YSLNVLVGLNDTMENLLAAVDRNEAEISPSITLAIACWENVPFINGSPQNTFVGLIDLA 295

Qy 296 IANTLLIGDDDFKSGQTKMKSVLDFPLVAGAKIPFISYSYHNLGNDOGAMLSAQTRRSK 355

Db 296 IANTLLIGDDDFKSGQTKMKSVLDFPLVAGAKIPFISYSYHNLGNDOGAMLSAQTRRSK 355

Qy 356 EISKSNVDDMNYSNALLYEPEHPDHVVLIKVPYVYDSSGRAMDEYTSSEIFPMGK 411

Db 356 EISKSNVDDMNYSNALLYEPEHPDHVVLIKVPYVYDSSGRAMDEYTSSEIFPMGK 411

```

RESULT 14
US-10-424-599-154863
: Sequence 154863, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 154863
: LENGTH: 431
: TYPE: PRT
: ORGANISM: Glycine max
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1.pep
: US-10-424-599-154863

```

Query Match	31.24	Score 159	DB 15	Length 431
Best Local Similarity	100.04	Prod. No. 1.2e-148		
Matches 159	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	236	YSNLVVGINDTMENTLLAADVDRNEAISPSTLYALACWENYEPFINGSPONTFVPLIDIA	295	
Db	236	YSNLVVGINDTMENTLLAADVDRNEAISPSTLYALACWENYEPFINGSPONTFVPLIDIA	295	
QY	296	IARNTLLIGGDDFKSGGQTMMKSVLVDPLVYAGAIKPTSIYSYHNLGNDDGNTSAPQTFRSK	355	
Db	296	IARNTLLIGGDDFKSGGQTMMKSVLVDPLVYAGAIKPTSIYSYHNLGNDDGNTSAPQTFRSK	355	
QY	356	EISKSNVVDMDVNSNALIYEPGEHDPDHVYVLYKYPYVYD	394	
Db	356	EISKSNVVDMDVNSNALIYEPGEHDPDHVYVLYKYPYVYD	394	

```

RESULT 15
US-10-424-599-259439
; Sequence 259439, Application US/10424599
; Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ. ID NOS: 285684
SEQ ID NO 259439
LENGTH: 220
TYPE: PRT

```

```

? ORGANISM: Glycine max
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (1)..(220)
? OTHER INFORMATION: unsure at all Xaa locations
? FEATURE:
? OTHER INFORMATION: Clone ID: PAT_MRT3847_76298C.1.pep
? OS-10-424-599-259439

```

Query Match 22.2%; Score 113; DB 15; Length 220;
 Best Local Similarity 100.0%; Pred. No. 2.9e-103;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY RNEAISPSTLYAACMENVPFNGSPONTFVPGILDIAARNTLIGDDPFKSGQTKK 315
 |||
 Db 2 RNEAISPSTLYAACMENVPFNGSPONTFVPGILDIAARNTLIGDDPFKSGQTKK 61

QY SVLVDFVGVGAIKPTISVSYNHLGNDGMLSAQTFRSKEISKSNVDDWVN 368
 |||
 Db 62 SVLVDFVGVGAIKPTISVSYNHLGNDGMLSAQTFRSKEISKSNVDDWVN 114

Search completed: June 7, 2005, 17:25:33
Job time : 78.5 secs

Job time : 78.5 secs

Job time : 78.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:03:16 ; Search time 276.5 Seconds

(without alignments)
2154.370 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MEIENPKVCSPVKYETETI.....NIMRACVGLAPNNMILEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 6959266 seqs, 116806243 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Pending Patents AA Main:*

```
1: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	510	100.0	510	12	US-08-835-751A-2	Sequence 2, Appl1
2	510	100.0	510	16	US-09-299-315-2	Sequence 2, Appl1
3	510	100.0	510	16	US-09-299-315-10	Sequence 10, Appl1
4	510	100.0	510	17	US-09-304-534-2	Sequence 2, Appl1
5	510	100.0	510	17	US-09-304-534-10	Sequence 10, Appl1
6	510	100.0	510	26	US-10-025-003-2	Sequence 2, Appl1
7	510	100.0	510	26	US-10-025-003-10	Sequence 10, Appl1
8	510	100.0	510	33	US-10-718-952-2	Sequence 2, Appl1
9	510	100.0	510	33	US-10-718-952-10	Sequence 10, Appl1
10	510	100.0	510	37	US-60-592-978-4180	Sequence 4180, Ap
11	510	100.0	510	37	US-60-592-978-6132	Sequence 6132, Ap
12	440	86.3	511	30	US-10-424-599-213009	Sequence 213009,
13	409	80.2	510	12	US-08-835-751A-6	Sequence 6, Appl1
14	409	80.2	510	16	US-09-299-315-6	Sequence 6, Appl1
15	409	80.2	510	17	US-09-304-534-6	Sequence 6, Appl1
16	409	80.2	510	26	US-10-025-003-6	Sequence 6, Appl1
17	409	80.2	510	33	US-10-718-952-6	Sequence 6, Appl1
18	376	73.7	510	37	US-60-592-978-13648	Sequence 13648, A
19	253	49.6	510	16	US-09-299-315-14	Sequence 14, Appl1
20	253	49.6	510	16	US-09-299-315-16	Sequence 16, Appl1
21	253	49.6	510	17	US-09-304-534-14	Sequence 14, Appl1
22	253	49.6	510	17	US-09-304-534-16	Sequence 16, Appl1
23	253	49.6	510	26	US-10-025-003-14	Sequence 14, Appl1
24	253	49.6	510	26	US-10-025-003-16	Sequence 16, Appl1
25	253	49.6	510	33	US-10-718-952-14	Sequence 14, Appl1
26	253	49.6	510	33	US-10-718-952-16	Sequence 16, Appl1
27	253	49.6	510	37	US-60-592-978-12850	Sequence 12850, A
28	236	46.3	510	17	US-09-299-315-12	Sequence 12, Appl1
29	236	46.3	510	16	US-09-304-534-12	Sequence 12, Appl1
30	236	46.3	510	26	US-10-025-003-12	Sequence 12, Appl1
31	236	46.3	510	33	US-10-718-952-12	Sequence 12, Appl1
32	165	32.4	636	37	US-60-592-978-15981	Sequence 15981, A
33	159	31.2	431	30	US-10-424-599-154863	Sequence 154863,
34	115	22.5	184	37	US-60-592-978-2477	Sequence 2477, Ap
35	113	22.2	220	30	US-10-424-599-259439	Sequence 259439,
36	113	22.2	220	37	US-60-592-978-10124	Sequence 10124, A
37	111	21.8	260	37	US-60-592-978-20424	Sequence 20424, A
38	110	21.6	148	37	US-60-592-978-7377	Sequence 7377, Ap
39	103	20.2	211	30	US-10-424-599-165505	Sequence 165505,
40	103	20.2	276	37	US-60-592-978-21322	Sequence 21322, A
41	102	20.0	510	30	US-10-424-599-154864	Sequence 154864,
42	97	19.0	211	37	US-60-592-978-4373	Sequence 4373, Ap
43	94	18.4	557	37	US-60-592-978-21798	Sequence 21798, A
44	92	18.0	472	37	US-60-592-978-4518	Sequence 4518, Ap
45	79	15.5	226	37	US-60-592-978-15017	Sequence 15017, A

ALIGNMENTS

```
RESULT 1
US-08-835-751A-2
; Sequence 2, Application US/08835751A
; GENERAL INFORMATION:
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: SEBASTIAN, SCOTT ANTHONY
; TITLE OF INVENTION: SOYBEAN PLANTS PRODUCING SEEDS WITH
; TITLE OF INVENTION: REDUCED LEVELS OF RAFFINOSE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
```

```

; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,751A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: MAJARIAN, WILLIAM R.
; REGISTRATION NUMBER: P41,173
; REFERENCE/DOCKET NUMBER: BB-1077
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302)992-4926
; TELEFAX: (302)773-0164
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-835-751A-2
```

```

Query Match          100.0%; Score 510; DB 12; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MFLENFKVCEPNVKYTEFTEIOSVYNYETTELVHENRNGTYOMIVKPKSVKTEFKTNIHVP 60
DB 1 MFLENFKVCEPNVKYTEFTEIOSVYNYETTELVHENRNGTYOMIVKPKSVKTEFKTNIHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPWNPDPDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDPDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTERYSNLV 240
QY 241 VGLNDTMENLLAADVDRNEAISPSTLYAIACWENVPTINGSPONTFVGLIDLAIAANT 300
DB 241 VGLNDTMENLLAADVDRNEAISPSTLYAIACWENVPTINGSPONTFVGLIDLAIAANT 300
QY 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
QY 361 NVYDDMVNSNALIYEPGHPDHVVVVKYVPYVGSKRAMDEYTSSEIFMGKSTIYLHNTC 420
DB 361 NVYDDMVNSNALIYEPGHPDHVVVVKYVPYVGSKRAMDEYTSSEIFMGKSTIYLHNTC 420
QY 421 EDSLAAPIIIDLVLLAEISTRIFEKANEKGKPHSFHVAITILSYLTAKPLVPGTPVNV 480
DB 421 EDSLAAPIIIDLVLLAEISTRIFEKANEKGKPHSFHVAITILSYLTAKPLVPGTPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

```

RESULT 2
US-09-299-315-2
; Sequence 2, Application US/09299315
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelet, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
```

```

; FILE REFERENCE: BB-1077-B
; CURRENT APPLICATION NUMBER: US/09/299,315
; CURRENT FILING DATE: 1999-04-26
; EARLIER APPLICATION NUMBER: 08/835,751
; EARLIER FILING DATE: APRIL 8, 1997
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-09-299-315-2
```

```

Query Match          100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MFLENFKVCEPNVKYTEFTEIOSVYNYETTELVHENRNGTYOMIVKPKSVKTEFKTNIHVP 60
DB 1 MFLENFKVCEPNVKYTEFTEIOSVYNYETTELVHENRNGTYOMIVKPKSVKTEFKTNIHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPWNPDPDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDPDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIKAFKATKVDKVVVLTANTERYSNLV 240
QY 241 VGLNDTMENLLAADVDRNEAISPSTLYAIACWENVPTINGSPONTFVGLIDLAIAANT 300
DB 241 VGLNDTMENLLAADVDRNEAISPSTLYAIACWENVPTINGSPONTFVGLIDLAIAANT 300
QY 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
QY 361 NVYDDMVNSNALIYEPGHPDHVVVVKYVPYVGSKRAMDEYTSSEIFMGKSTIYLHNTC 420
DB 361 NVYDDMVNSNALIYEPGHPDHVVVVKYVPYVGSKRAMDEYTSSEIFMGKSTIYLHNTC 420
QY 421 EDSLAAPIIIDLVLLAEISTRIFEKANEKGKPHSFHVAITILSYLTAKPLVPGTPVNV 480
DB 421 EDSLAAPIIIDLVLLAEISTRIFEKANEKGKPHSFHVAITILSYLTAKPLVPGTPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

```

RESULT 3
US-09-299-315-10
; Sequence 10, Application US/09299315
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelet, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-B
; CURRENT APPLICATION NUMBER: US/09/299,315
; CURRENT FILING DATE: 1999-04-26
; EARLIER APPLICATION NUMBER: 08/835,751
; EARLIER FILING DATE: APRIL 8, 1997
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
```

TYPE: PRT
ORGANISM: Glycine max
US-09-299-315-10

Query Match 100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MFIENFKVCEPNVKYETEIEIOSVYNYETTELVEHNRNGTYQWIVKPKSVKYEKKNIHVP 60
DB 1 MFIENFKVCEPNVKYETEIEIOSVYNYETTELVEHNRNGTYQWIVKPKSVKYEKKNIHVP 60
QY 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPMVPDDIVFGWMDISNMNLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPMVPDDIVFGWMDISNMNLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
QY 181 YPDPFIAANOEBRANNVIGTKOEOVQOIIKDKAFKATKVDKVVVLTANTERSNLV 240
DB 181 YPDPFIAANOEBRANNVIGTKOEOVQOIIKDKAFKATKVDKVVVLTANTERSNLV 240
QY 241 VGLNDTMENLAAVDREAEISPTLYAIAQWENVPFNGSPONTFVGLIDLAIRMT 300
DB 241 VGLNDTMENLAAVDREAEISPTLYAIAQWENVPFNGSPONTFVGLIDLAIRMT 300
QY 301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360
DB 301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360
QY 361 NVYDDMVNSNALIYEBGEHPDHVVVIKYVPYVGDSCRADDEYSEIFMGKSTIVLHNTC 420
DB 361 NVYDDMVNSNALIYEBGEHPDHVVVIKYVPYVGDSCRADDEYSEIFMGKSTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIEFKAENEGFHSFHPVATILSYLTKAPLVPGTPVNV 480
DB 421 EDSLAAPIIIDLVLAELSTRIEFKAENEGFHSFHPVATILSYLTKAPLVPGTPVNV 480
QY 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510

```

RESULT 4
US-09-304-534-2
Sequence 2, Application US/09304534
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebaastian, Scott
APPLICANT: Grace, John
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/09/304, 534
EARLIER FILING DATE: 1999-05-04
EARLIER APPLICATION NUMBER: 08/835, 751
EARLIER FILING DATE: APRIL 8, 1997
EARLIER APPLICATION NUMBER: PCT/US98/06822
EARLIER FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-09-304-534-2

Query Match 100.0%; Score 510; DB 17; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MFIENFKVCEPNVKYETEIEIOSVYNYETTELVEHNRNGTYQWIVKPKSVKYEKKNIHVP 60
DB 1 MFIENFKVCEPNVKYETEIEIOSVYNYETTELVEHNRNGTYQWIVKPKSVKYEKKNIHVP 60
QY 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPMVPDDIVFGWMDISNMNLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPMVPDDIVFGWMDISNMNLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
QY 181 YPDPFIAANOEBRANNVIGTKOEOVQOIIKDKAFKATKVDKVVVLTANTERSNLV 240
DB 181 YPDPFIAANOEBRANNVIGTKOEOVQOIIKDKAFKATKVDKVVVLTANTERSNLV 240
QY 241 VGLNDTMENLAAVDREAEISPTLYAIAQWENVPFNGSPONTFVGLIDLAIRMT 300
DB 241 VGLNDTMENLAAVDREAEISPTLYAIAQWENVPFNGSPONTFVGLIDLAIRMT 300
QY 301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360
DB 301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360
QY 361 NVYDDMVNSNALIYEBGEHPDHVVVIKYVPYVGDSCRADDEYSEIFMGKSTIVLHNTC 420
DB 361 NVYDDMVNSNALIYEBGEHPDHVVVIKYVPYVGDSCRADDEYSEIFMGKSTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIEFKAENEGFHSFHPVATILSYLTKAPLVPGTPVNV 480
DB 421 EDSLAAPIIIDLVLAELSTRIEFKAENEGFHSFHPVATILSYLTKAPLVPGTPVNV 480
QY 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510

```

RESULT 5
US-09-304-534-10
Sequence 10, Application US/09304534
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebaastian, Scott
APPLICANT: Grace, John
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/09/304, 534
EARLIER FILING DATE: 1999-05-04
EARLIER APPLICATION NUMBER: 08/835, 751
EARLIER FILING DATE: APRIL 8, 1997
EARLIER APPLICATION NUMBER: PCT/US98/06822
EARLIER FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 10
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-09-304-534-10

Query Match 100.0%; Score 510; DB 17; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MFIENFKVCEPNVKYETEIEIOSVYNYETTELVEHNRNGTYQWIVKPKSVKYEKKNIHVP 60
DB 1 MFIENFKVCEPNVKYETEIEIOSVYNYETTELVEHNRNGTYQWIVKPKSVKYEKKNIHVP 60
QY 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

```

```
Db      61 KLGWLVGWMGNGNSTLTGGVIANREGISWATDKIQOANFVGSITQASAIRVGSFOGEE 120
Qy      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKAFVDIDLOKOLRPYMESMLPLPGI 180
Db      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKAFVDIDLOKOLRPYMESMLPLPGI 180
Qy      181 YDPDFIAANOEBERANNAVIGTKQEOVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Db      181 YDPDFIAANOEBERANNAVIGTKQEOVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Qy      241 VGLNTMTENLLAAVDRNEAISPSTLYAIACMENVPIFNGSPONTFVPGILDLAIARNT 300
Db      241 VGLNTMTENLLAAVDRNEAISPSTLYAIACMENVPIFNGSPONTFVPGILDLAIARNT 300
Qy      301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
Db      301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
Qy      361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db      361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy      421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFSFHPVATILSYLTAKPLVPFGTPPVN 480
Db      421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFSFHPVATILSYLTAKPLVPFGTPPVN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
```

```
RESULT 6
US-10-025-003-2
; Sequence 2, Application US/10025003
```

```
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-2
```

```
Query Match      100.0%; Score 510; DB 26; Length 510;
```

```
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
```

```
Matches 510; Conservative 0; Mismatches 0;
```

```
Qy      1 MEIENFKVCEPNVKYKTEIEIOSVYNYETTELVEHNRNGTYOWIVPKSVKYEFTKNIHVP 60
Db      1 MEIENFKVCEPNVKYKTEIEIOSVYNYETTELVEHNRNGTYOWIVPKSVKYEFTKNIHVP 60
Qy      61 KLGWLVGWMGNGNSTLTGGVIANREGISWATDKIQOANFVGSITQASAIRVGSFOGEE 120
Db      61 KLGWLVGWMGNGNSTLTGGVIANREGISWATDKIQOANFVGSITQASAIRVGSFOGEE 120
Qy      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKAFVDIDLOKOLRPYMESMLPLPGI 180
Db      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKAFVDIDLOKOLRPYMESMLPLPGI 180
Qy      181 YDPDFIAANOEBERANNAVIGTKQEOVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Db      181 YDPDFIAANOEBERANNAVIGTKQEOVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
```

```
Db      181 YDPDFIAANOEBERANNAVIGTKQEOVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Qy      241 VGLNTMTENLLAAVDRNEAISPSTLYAIACMENVPIFNGSPONTFVPGILDLAIARNT 300
Db      241 VGLNTMTENLLAAVDRNEAISPSTLYAIACMENVPIFNGSPONTFVPGILDLAIARNT 300
Qy      301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
Db      301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
Qy      361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db      361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy      421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFSFHPVATILSYLTAKPLVPFGTPPVN 480
Db      421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFSFHPVATILSYLTAKPLVPFGTPPVN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
```

```
RESULT 7
US-10-025-003-10
; Sequence 10, Application US/10025003
```

```
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-10
```

```
Query Match      100.0%; Score 510; DB 26; Length 510;
```

```
Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
```

```
Matches 510; Conservative 0; Mismatches 0;
```

```
Qy      1 MEIENFKVCEPNVKYKTEIEIOSVYNYETTELVEHNRNGTYOWIVPKSVKYEFTKNIHVP 60
Db      1 MEIENFKVCEPNVKYKTEIEIOSVYNYETTELVEHNRNGTYOWIVPKSVKYEFTKNIHVP 60
Qy      61 KLGWLVGWMGNGNSTLTGGVIANREGISWATDKIQOANFVGSITQASAIRVGSFOGEE 120
Db      61 KLGWLVGWMGNGNSTLTGGVIANREGISWATDKIQOANFVGSITQASAIRVGSFOGEE 120
Qy      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKAFVDIDLOKOLRPYMESMLPLPGI 180
Db      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKAFVDIDLOKOLRPYMESMLPLPGI 180
Qy      181 YDPDFIAANOEBERANNAVIGTKQEOVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Db      181 YDPDFIAANOEBERANNAVIGTKQEOVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Qy      241 VGLNTMTENLLAAVDRNEAISPSTLYAIACMENVPIFNGSPONTFVPGILDLAIARNT 300
Db      241 VGLNTMTENLLAAVDRNEAISPSTLYAIACMENVPIFNGSPONTFVPGILDLAIARNT 300
```



```
QY 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNISAPOTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNISAPOTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDLSLAAPITLLDVLIAELSTRIEFKANEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Db 421 EDLSLAAPITLLDVLIAELSTRIEFKANEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
QY 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

RESULT 8
US-10-718-952-2
; Sequence 2, Application US/10718952
; GENERAL INFORMATION:
; APPLICANT: Hietz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-2

Query Match      100.0%; Score 510; DB 33; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFIEENFKVBCPNVKYKTEIEIQSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Db 1 MFIEENFKVBCPNVKYKTEIEIQSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
QY 61 KIGWMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
Db 61 KIGWMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFSSLLPMVNPDIYFGGMDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFSSLLPMVNPDIYFGGMDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIFKFEATKYDKVVLVMTANTERSNLV 240
Db 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIFKFEATKYDKVVLVMTANTERSNLV 240
QY 241 VGLNDTMENTLLAADRNEAISPSTLYAIACWENVPFINGSPONTFVGLIDLAIARNT 300
Db 241 VGLNDTMENTLLAADRNEAISPSTLYAIACWENVPFINGSPONTFVGLIDLAIARNT 300
QY 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNISAPOTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNISAPOTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGKSTIVLHNTC 420
```

```
QY 421 EDLSLAAPITLLDVLIAELSTRIEFKANEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Db 421 EDLSLAAPITLLDVLIAELSTRIEFKANEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
QY 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

RESULT 9
US-10-718-952-10
; Sequence 10, Application US/10718952
; GENERAL INFORMATION:
; APPLICANT: Hietz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-10

Query Match      100.0%; Score 510; DB 33; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFIEENFKVBCPNVKYKTEIEIQSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Db 1 MFIEENFKVBCPNVKYKTEIEIQSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
QY 61 KIGWMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
Db 61 KIGWMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFSSLLPMVNPDIYFGGMDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFSSLLPMVNPDIYFGGMDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIFKFEATKYDKVVLVMTANTERSNLV 240
Db 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIFKFEATKYDKVVLVMTANTERSNLV 240
QY 241 VGLNDTMENTLLAADRNEAISPSTLYAIACWENVPFINGSPONTFVGLIDLAIARNT 300
Db 241 VGLNDTMENTLLAADRNEAISPSTLYAIACWENVPFINGSPONTFVGLIDLAIARNT 300
QY 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNISAPOTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNISAPOTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDLSLAAPITLLDVLIAELSTRIEFKANEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Db 421 EDLSLAAPITLLDVLIAELSTRIEFKANEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
QY 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
```

Db 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

RESULT 10

US-60-592-978-4180

Sequence 4180, Application US/60592978

GENERAL INFORMATION:

APPLICANT: Abad, Mark S.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53452)A

CURRENT APPLICATION NUMBER: US/60/592,978

CURRENT FILING DATE: 2004-07-31

NUMBER OF SEQ ID NOS: 22568

SEQ ID NO 4180

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-60-592-978-4180

Query Match 100.0%; Score 510; DB 37; Length 510;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENFKVECPVVKYTEIEIOSVYNETTELVEHNRNGTYOMIVKPKSVKYEFTKNIHVP 60

Db 1 MEIENFKVECPVVKYTEIEIOSVYNETTELVEHNRNGTYOMIVKPKSVKYEFTKNIHVP 60

QY 61 KLGWMLVGMGNGNSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVGSFQGBE 120

Db 61 KLGWMLVGMGNGNSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVGSFQGBE 120

QY 121 IYAPFKSLPMPVNPDPDIFGGMDISMNLADAMARAKVFDIDLOKOLAPYMESMLPLPGI 180

Db 121 IYAPFKSLPMPVNPDPDIFGGMDISMNLADAMARAKVFDIDLOKOLAPYMESMLPLPGI 180

QY 181 YDPDFIAANOEBRANNVIKGTQOEVOQIINDIKAFKATKVDKVVLTANTERTYSNLV 240

Db 181 YDPDFIAANOEBRANNVIKGTQOEVOQIINDIKAFKATKVDKVVLTANTERTYSNLV 240

QY 241 VGLNDTMENLLAAVDNRNAEISPTLYAIACMENVPIFGSPONTFVGLIDLAIAANT 300

Db 241 VGLNDTMENLLAAVDNRNAEISPTLYAIACMENVPIFGSPONTFVGLIDLAIAANT 300

QY 301 LIGGDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360

Db 301 LIGGDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360

QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

QY 421 EDSLAAPIILDLVLAELSTRIEFKANEKGKFSHPVATILSYLTKAPLVPPTPVVN 480

Db 421 EDSLAAPIILDLVLAELSTRIEFKANEKGKFSHPVATILSYLTKAPLVPPTPVVN 480

QY 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

Db 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

RESULT 11

US-60-592-978-6132

Sequence 6132, Application US/60592978

GENERAL INFORMATION:

APPLICANT: Abad, Mark S.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53452)A

CURRENT APPLICATION NUMBER: US/60/592,978

CURRENT FILING DATE: 2004-07-31

NUMBER OF SEQ ID NOS: 22568

SEQ ID NO 6132

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-60-592-978-6132

Query Match 100.0%; Score 510; DB 37; Length 510;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENFKVECPVVKYTEIEIOSVYNETTELVEHNRNGTYOMIVKPKSVKYEFTKNIHVP 60

Db 1 MEIENFKVECPVVKYTEIEIOSVYNETTELVEHNRNGTYOMIVKPKSVKYEFTKNIHVP 60

QY 61 KLGWMLVGMGNGNSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVGSFQGBE 120

Db 61 KLGWMLVGMGNGNSTLTGCVIANREGISWATKDKIQOANFGSLTQASAIRVGSFQGBE 120

QY 121 IYAPFKSLPMPVNPDPDIFGGMDISMNLADAMARAKVFDIDLOKOLAPYMESMLPLPGI 180

Db 121 IYAPFKSLPMPVNPDPDIFGGMDISMNLADAMARAKVFDIDLOKOLAPYMESMLPLPGI 180

QY 181 YDPDFIAANOEBRANNVIKGTQOEVOQIINDIKAFKATKVDKVVLTANTERTYSNLV 240

Db 181 YDPDFIAANOEBRANNVIKGTQOEVOQIINDIKAFKATKVDKVVLTANTERTYSNLV 240

QY 241 VGLNDTMENLLAAVDNRNAEISPTLYAIACMENVPIFGSPONTFVGLIDLAIAANT 300

Db 241 VGLNDTMENLLAAVDNRNAEISPTLYAIACMENVPIFGSPONTFVGLIDLAIAANT 300

QY 301 LIGGDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360

Db 301 LIGGDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360

QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

QY 421 EDSLAAPIILDLVLAELSTRIEFKANEKGKFSHPVATILSYLTKAPLVPPTPVVN 480

Db 421 EDSLAAPIILDLVLAELSTRIEFKANEKGKFSHPVATILSYLTKAPLVPPTPVVN 480

QY 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

Db 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

RESULT 12

US-10-424-599-213009

Sequence 213009, Application US/10424599

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 213009

LENGTH: 511

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(511)

OTHER INFORMATION: unsure at all Xaa locations

OTHER INFORMATION: Clone ID: PAT_MKT3847_34372C.1.dep

US-10-424-599-213009

Query Match 86.3%; Score 440; DB 30; Length 511;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 71 GNGSLTGTGVIANREGISWATKIOQANYPGSLTOASAIRVSGFGEETVAFKSLP 130
DB 72 GNGSLTGTGVIANREGISWATKIOQANYPGSLTOASAIRVSGFGEETVAFKSLP 131
QY 131 MNPDDIVFGWMDISNNMLADAMARAKVPDIDLOKOLRPMESMLPLPGIYDDPFIANO 190
DB 132 MNPDDIVFGWMDISNNMLADAMARAKVPDIDLOKOLRPMESMLPLPGIYDDPFIANO 191
QY 191 BEBANNVKTQBOVOQIIKDIAKFAKATKVDKVVLTANTERSNLVGLNDTMENT 250
DB 192 BEBANNVKTQBOVOQIIKDIAKFAKATKVDKVVLTANTERSNLVGLNDTMENT 251
QY 251 LAAVDNEARISPTLYAICWENVPFINGSPONTFVPGILDIAIARNTLIGDDPKSG 310
DB 252 LAAVDNEARISPTLYAICWENVPFINGSPONTFVPGILDIAIARNTLIGDDPKSG 311
QY 311 QTKKSVLVDPLVAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISKSNVDDMVNSN 370
DB 312 QTKKSVLVDPLVAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISKSNVDDMVNSN 371
QY 371 ALIYEGEHPDHVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIYLHNTCEDSLAAPIT 430
DB 372 ALIYEGEHPDHVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIYLHNTCEDSLAAPIT 431
QY 431 LDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPGTPVNNALSKORAMLE 490
DB 432 LDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPGTPVNNALSKORAMLE 491
QY 491 NIMRACVGLAPENNMLEYK 510
DB 492 NIMRACVGLAPENNMLEYK 511

```

RESULT 13

US-08-835-751A-6
; Sequence 6, Application US/08835751A

GENERAL INFORMATION:

APPLICANT: HITZ, WILLIAM D.

TITLE OF INVENTION: SOYBEAN PLANTS PRODUCING SEEDS WITH

TITLE OF INVENTION: REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

STREET: 1007 MARKET STREET

CITY: WILMINGTON

STATE: DELAWARE

COUNTRY: UNITED STATES OF AMERICA

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.50 INCH

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: MICROSOFT WINDOWS 95

SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (7.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/835, 751A

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: MAJARIAN, WILLIAM R.

REGISTRATION NUMBER: P41,173

REFERENCE/DOCKET NUMBER: BB-1077

TELECOMMUNICATION INFORMATION:

TELEPHONE: (302)992-4926

TELEFAX: (302)773-0164

SEQUENCE CHARACTERISTICS:

LENGTH: 510 amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-835-751A-6

Query Match 80.2%; Score 409; DB 12; Length 510;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MFLENKVECPNKKYETETIQSVYNYETTELHNNNGTYQWIVKPSVYKEKTNHVP 60
DB 1 MFLENKVECPNKKYETETIQSVYNYETTELHNNNGTYQWIVKPSVYKEKTNHVP 60
QY 61 KLGVMVLVGMGNGSLTGTGVIANREGISWATKIOQANYPGSLTOASAIRVSGFGE 120
DB 61 KLGVMVLVGMGNGSLTGTGVIANREGISWATKIOQANYPGSLTOASAIRVSGFGE 120
QY 121 IYAPFKSLPMPNPDDIVFGWMDISNNMLADAMARAKVPDIDLOKOLRPMESMLPLPGI 180
DB 121 IYAPFKSLPMPNPDDIVFGWMDISNNMLADAMARAKVPDIDLOKOLRPMESMLPLPGI 180
QY 181 YDDPFIANOEBEANNVYIKGTQBOVOQIIKDIAKFAKATKVDKVVLTANTERSNLV 240
DB 181 YDDPFIANOEBEANNVYIKGTQBOVOQIIKDIAKFAKATKVDKVVLTANTERSNLV 240
QY 241 VGLNDTMENTLAAVDNEARISPTLYAICWENVPFINGSPONTFVPGILDIAIARNT 300
DB 241 VGLNDTMENTLAAVDNEARISPTLYAICWENVPFINGSPONTFVPGILDIAIARNT 300
QY 301 LIGDDPKSGQTKKSVLVDPLVAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISK 360
DB 301 LIGDDPKSGQTKKSVLVDPLVAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISK 360
QY 361 NVYDDMVNSNALIYEGEHPDHVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIYLHNTC 420
DB 361 NVYDDMVNSNALIYEGEHPDHVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIYLHNTC 420
QY 421 EDSLAPITLDDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPGTPVNN 480
DB 421 EDSLAPITLDDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPGTPVNN 480
QY 481 ALSKORAMLENNIMRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENNIMRACVGLAPENNMLEYK 510

```

RESULT 14

US-09-299-315-6

; Sequence 6, Application US/09299315

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-B

CURRENT APPLICATION NUMBER: US/09/299, 315

EARLIER FILING DATE: 1999-04-26

EARLIER APPLICATION NUMBER: 08/835,751

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97

SEQ ID NO 6

LENGTH: 510

ORGANISM: Glycine max

US-09-299-315-6

```

Query Match 80.2%; Score 409; DB 16; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFLENKVECPNKKYETETIQSVYNYETTELHNNNGTYQWIVKPSVYKEKTNHVP 60

```

```

Db      1 MEIENFKVCEPNVKTETEIQSVYNYETTELVEHNNNGTYOWI VPKSKVYEFKTNIHVP 60
Qy      61 KGVNMLVGVGNNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGBE 120
Db      61 KGVNMLVGVGNNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGBE 120
Qy      121 IYAPFKSLPVMNPPDDIVFGWDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGI 180
Db      121 IYAPFKSLPVMNPPDDIVFGWDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGI 180
Qy      181 YPDPFIANAQOEERANNVIGTKOEVOQOIKDIAKFEATKVDKVVVLTANTERSNLY 240
Db      181 YPDPFIANAQOEERANNVIGTKOEVOQOIKDIAKFEATKVDKVVVLTANTERSNLY 240
Qy      241 VGLNDTMENLLAAVDRENEAISPSTLYAIACMENVPFINGSPONTFVPGGLDLAIARNT 300
Db      241 VGLNDTMENLLAAVDRENEAISPSTLYAIACMENVPFINGSPONTFVPGGLDLAIARNT 300
Qy      301 LIGDDFFKSGQTKMSVLDVFLVGAGIKPTSIIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
Db      301 LIGDDFFKSGQTKMSVLDVFLVGAGIKPTSIIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
Qy      361 NVDDMVNSNALLYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db      361 NVDDMVNSNALLYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Qy      421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFHSPVATILSYLTAKAPLVPPTPVVN 480
Db      421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFHSPVATILSYLTAKAPLVPPTPVVN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

```

RESULT 15
US-09-304-534-6
; Sequence 6, Application US/09304534
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebaetian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelt, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/09/304,534
; CURRENT FILING DATE: 1999-05-04
; EARLIER APPLICATION NUMBER: 08/835,751
; EARLIER FILING DATE: APRIL 8, 1997
; EARLIER APPLICATION NUMBER: PCT/US98/06822
; EARLIER FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-09-304-534-6

```

```

Query Match      80.2%; Score 409; DB 17; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 MEIENFKVCEPNVKTETEIQSVYNYETTELVEHNNNGTYOWI VPKSKVYEFKTNIHVP 60
Db      1 MEIENFKVCEPNVKTETEIQSVYNYETTELVEHNNNGTYOWI VPKSKVYEFKTNIHVP 60
Qy      61 KGVNMLVGVGNNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGBE 120
Db      61 KGVNMLVGVGNNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGBE 120
Qy      121 IYAPFKSLPVMNPPDDIVFGWDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGI 180

```

```

Db      121 IYAPFKSLPVMNPPDDIVFGWDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGI 180
Qy      181 YPDPFIANAQOEERANNVIGTKOEVOQOIKDIAKFEATKVDKVVVLTANTERSNLY 240
Db      181 YPDPFIANAQOEERANNVIGTKOEVOQOIKDIAKFEATKVDKVVVLTANTERSNLY 240
Qy      241 VGLNDTMENLLAAVDRENEAISPSTLYAIACMENVPFINGSPONTFVPGGLDLAIARNT 300
Db      241 VGLNDTMENLLAAVDRENEAISPSTLYAIACMENVPFINGSPONTFVPGGLDLAIARNT 300
Qy      301 LIGDDFFKSGQTKMSVLDVFLVGAGIKPTSIIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
Db      301 LIGDDFFKSGQTKMSVLDVFLVGAGIKPTSIIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
Qy      361 NVDDMVNSNALLYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db      361 NVDDMVNSNALLYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Qy      421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFHSPVATILSYLTAKAPLVPPTPVVN 480
Db      421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFHSPVATILSYLTAKAPLVPPTPVVN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

```

Search completed: June 7, 2005, 17:21:12
Job time : 278.5 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 7, 2005, 17:06:41 ; Search time 41.5 Seconds
(without alignments)
1462.923 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MFIENPKVCEPCVKKYETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 557933 seqs, 119041758 residues

Word size : 0

Total number of hits satisfying chosen parameters: 557933

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database :

Pending Patents AA New:*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US11_NEW_COMB.pep:*
8: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	14.7	513	6	US-10-703-032-108934, Sequence 108934, A
2	67	13.1	192	6	US-10-703-032-108933, Sequence 108933, A
3	67	13.1	510	6	US-10-703-032-135209, Sequence 135209, A
4	67	13.1	510	7	US-11-064-295-4, Sequence 4, App11
5	29	5.7	565	7	US-11-097-143-23283, Sequence 23283, Ap
6	19	3.7	235	6	US-10-990-328A-8846, Sequence 8846, Ap
7	18	3.5	132	6	US-10-703-032-147024, Sequence 147024, A
8	17	3.3	111	6	US-10-450-763-34728, Sequence 34728, A
9	17	3.3	280	8	US-60-669-175-37019, Sequence 37019, A
10	17	3.3	558	6	US-10-990-328A-10375, Sequence 10375, A
11	17	3.3	559	8	US-60-671-501-362, Sequence 362, App
12	15	2.9	704	6	US-10-450-763-52092, Sequence 52092, A
13	15	2.9	704	6	US-10-450-763-34498, Sequence 34498, A
14	13	2.5	244	6	US-10-450-763-34497, Sequence 34497, A
15	13	2.5	254	6	US-10-450-763-34497, Sequence 34497, A
16	13	2.5	266	6	US-10-450-763-52091, A Sequence 52091, A
17	13	2.5	533	6	US-10-989-488A-73, Sequence 10374, A
18	13	2.5	534	6	US-10-989-488A-73, Sequence 73, App1
19	11	2.2	84	6	US-10-450-763-34729, Sequence 34729, A
20	10	2.0	116	6	US-10-703-032-204923, Sequence 204923, A
21	10	2.0	142	6	US-10-450-763-34727, Sequence 34727, A
22	8	1.6	113	6	US-10-863-332A-196, Sequence 196, App
23	8	1.6	378	8	US-60-643-717-5228, Sequence 5228, Ap
24	8	1.6	408	8	US-60-643-717-10051, Sequence 10051, A
25	7	1.4	90	6	US-10-703-032-154073, Sequence 154073, A

26	7	1.4	93	7	US-11-031-175-15484, Sequence 15484, A
27	7	1.4	103	6	US-10-703-032-189322, Sequence 189322, A
28	7	1.4	123	7	US-11-082-389-110, Sequence 110, App
29	7	1.4	126	6	US-10-703-032-149462, Sequence 149462, A
30	7	1.4	126	6	US-10-703-032-177833, Sequence 177833, A
31	7	1.4	127	6	US-10-450-763-59912, Sequence 59912, A
32	7	1.4	144	6	US-10-703-032-186174, Sequence 186174, A
33	7	1.4	151	7	US-11-082-389-108, Sequence 108, App
34	7	1.4	159	6	US-10-450-763-35496, Sequence 35496, A
35	7	1.4	161	6	US-10-703-032-126314, Sequence 126314, A
36	7	1.4	163	6	US-10-703-032-158910, Sequence 158910, A
37	7	1.4	175	8	US-60-655-875-147137, Sequence 147137, A
38	7	1.4	181	6	US-10-703-032-134690, Sequence 134690, A
39	7	1.4	182	6	US-10-703-032-210750, Sequence 210750, A
40	7	1.4	186	8	US-60-669-175-24108, Sequence 24108, A
41	7	1.4	188	7	US-11-079-463-9124, Sequence 9124, App
42	7	1.4	188	7	US-11-079-045-9124, Sequence 9124, App
43	7	1.4	192	6	US-10-703-032-156958, Sequence 156958, A
44	7	1.4	209	6	US-10-703-032-184953, Sequence 184953, A
45	7	1.4	215	6	US-10-703-032-174722, Sequence 174722, A

ALIGNMENTS

```
RESULT 1
US-10-703-032-108934
; Sequence 108934, Application US/10703032
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Cao, Yongwei
; APPLICANT: Masucci, James D.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21 (53374)B
; CURRENT APPLICATION NUMBER: US/10/703,032
; PRIOR APPLICATION NUMBER: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/020,338
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 21164
; SEQ ID NO 108934
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Trifolium aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_TA_3352.pep
; US-10-703-032-108934

Query Match          14.7%; Score 75; DB 6; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.8e-68;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIIYEPEDHDVVIKYVYVGDSCRAMEYTSSEIFMGKSTIVLANTCEDSLAAP 428
   |||||
DB 369 SNAIIYEPEDHDVVIKYVYVGDSCRAMEYTSSEIFMGKSTIVLANTCEDSLAAP 428
   |||||
QY 429 IILDVLAELSTRI 443
   |||||
DB 429 IILDVLAELSTRI 443
   |||||

RESULT 2
US-10-703-032-108933
; Sequence 108933, Application US/10703032
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Conner, Timothy W.
```

```
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Masucci, James D.
/ APPLICANT: Zhou, Yihua
/ TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53374)B
/ CURRENT APPLICATION NUMBER: US/10/703,032
/ CURRENT FILING DATE: 2003-11-06
/ PRIOR APPLICATION NUMBER: 10/020,338
/ PRIOR FILING DATE: 2001-12-12
/ NUMBER OF SEQ ID NOS: 21164
/ SEQ ID NO 108933
/ LENGTH: 192
/ TYPE: PRT
/ ORGANISM: Triticum aestivum
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_TA_3351.pep
US-10-703-032-108933
```

```
Query Match      13.1%; Score 67; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.2e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      301 LIGGDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
DB      99 LIGGDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 158
```

```
QY      361 NVVDDMV 367
DB      159 NVVDDMV 165
```

RESULT 3

```
US-10-703-032-135209
/ Sequence 135209, Application US/10703032
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Andersen, Scott E.
/ APPLICANT: Byrum, Joseph R.
/ APPLICANT: Comer, Timothy W.
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Masucci, James D.
/ APPLICANT: Zhou, Yihua
/ TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53374)B
/ CURRENT APPLICATION NUMBER: US/10/703,032
/ CURRENT FILING DATE: 2003-11-06
/ PRIOR APPLICATION NUMBER: 10/020,338
/ PRIOR FILING DATE: 2001-12-12
/ NUMBER OF SEQ ID NOS: 21164
/ SEQ ID NO 135209
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Triticum aestivum
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_TA_29627.pep
US-10-703-032-135209
```

```
Query Match      13.1%; Score 67; DB 6; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.2e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      301 LIGGDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
DB      301 LIGGDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
```

```
QY      361 NVVDDMV 367
DB      361 NVVDDMV 367
```

RESULT 4

```
US-11-064-295-4
/ Sequence 4, Application US/11064295
/ GENERAL INFORMATION:
/ APPLICANT: Liebergessel, Mathias
/ APPLICANT: Shi, Jintui
/ APPLICANT: Singletary, George
/ TITLE OF INVENTION: MODULATING MYO-INOSITOL CATABOLISM IN
/ TITLE OF INVENTION: PLANTS
/ FILE REFERENCE: 35718/287607
/ CURRENT APPLICATION NUMBER: US/11/064,295
/ CURRENT FILING DATE: 2005-02-23
/ PRIOR APPLICATION NUMBER: 60/547,640
/ PRIOR FILING DATE: 2004-02-25
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Zea mays
US-11-064-295-4
```

```
Query Match      13.1%; Score 67; DB 7; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.2e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      301 LIGGDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
DB      301 LIGGDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
```

```
QY      361 NVVDDMV 367
DB      361 NVVDDMV 367
```

RESULT 5

```
US-11-097-143-23283
/ Sequence 23283, Application US/11097143
/ GENERAL INFORMATION:
/ APPLICANT: Venier, J. Craig
/ APPLICANT: et al.
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 23283
/ LENGTH: 565
/ TYPE: PRT
/ ORGANISM: DROSOPHILA
US-11-097-143-23283
```

```
Query Match      5.7%; Score 29; DB 7; Length 565;
Best Local Similarity 100.0%; Pred. No. 5.7e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 378 EHPDHVVVTKVYPVYVGSXKRAMDEYTSI 406
DB 379 EHPDHVVVTKVYPVYVGSXKRAMDEYTSI 407

RESULT 6

US-10-990-328A-8846
Sequence 8846, Application US/10990328A
GENERAL INFORMATION:
APPLICANT: CAGGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
FILE REFERENCE: C1001495
CURRENT APPLICATION NUMBER: US/10/990,328A
CURRENT FILING DATE: 2004-11-17
NUMBER OF SEQ ID NOS: 558824
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8846
LENGTH: 235
TYPE: PRT
ORGANISM: Homo sapiens
US-10-990-328A-8846

Query Match 3.7%; Score 19; DB 6; Length 235;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 VLNHTCEDSLAAPILIDL 433
DB 50 VLNHTCEDSLAAPILIDL 68

RESULT 7

US-10-703-032-147024
Sequence 147024, Application US/10703032
GENERAL INFORMATION:
APPLICANT: Kovacic, David K.
APPLICANT: Andersen, Scott E.
APPLICANT: Byrum, Joseph R.
APPLICANT: Conner, Timothy W.
APPLICANT: Cao, Yongwei
APPLICANT: Masucci, James D.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 38-21(53374)B
CURRENT APPLICATION NUMBER: US/10/703,032
CURRENT FILING DATE: 2003-11-06
PRIOR APPLICATION NUMBER: 10/020,338
PRIOR FILING DATE: 2001-12-12
NUMBER OF SEQ ID NOS: 211164
SEQ ID NO 147024
LENGTH: 132
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: PAT_TA_41442.pcp
US-10-703-032-147024

Query Match 3.5%; Score 18; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 KVDKVVLTANTERYSN 238
DB 105 KVDKVVLTANTERYSN 122

RESULT 8
US-10-450-763-34728
Sequence 34728, Application US/10450763

GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US

CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 34728
LENGTH: 111
TYPE: PRT
ORGANISM: Homo sapiens
US-10-450-763-34728

Query Match 3.3%; Score 17; DB 6; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 LGVMLVGMGNGNSTLT 78
DB 76 LGVMLVGMGNGNSTLT 92

RESULT 9

US-60-669-175-37019
Sequence 37019, Application US/60669175
GENERAL INFORMATION:
APPLICANT: Monsanto Technology LLC
APPLICANT: Baum, James A
APPLICANT: Gilbertson, Larry A
APPLICANT: Kovacic, David K
APPLICANT: Larosa, Thomas J
APPLICANT: Lu, Maolong
APPLICANT: Munyikwa, Tichifa R. I.
APPLICANT: Roberts, James K
APPLICANT: Wu, Wei
APPLICANT: Zhang, Bei
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CONTROL OF INSECT INFESTATION IN PLANT
FILE REFERENCE: 38-21(53597)
CURRENT APPLICATION NUMBER: US/60/669,175
CURRENT FILING DATE: 2005-04-07
PRIOR APPLICATION NUMBER: 60560842
PRIOR FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 60565632
PRIOR FILING DATE: 2004-04-27
PRIOR APPLICATION NUMBER: 60579062
PRIOR FILING DATE: 2004-06-11
PRIOR APPLICATION NUMBER: 60603421
PRIOR FILING DATE: 2004-08-20
PRIOR APPLICATION NUMBER: 60617261
PRIOR FILING DATE: 2004-10-11
NUMBER OF SEQ ID NOS: 40774
SEQ ID NO 37019
LENGTH: 280
TYPE: PRT
ORGANISM: Diabrotica virgifera
FEATURE:
NAME/KEY: misc feature
LOCATION: (239)..(239)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_16716; Strand=+; Position=1-4
OTHER INFORMATION: Homolog annotation: Hit ID=XP_214319.2; Match level="QueryCoverage=99%, HitCoverage=50%, E-value=9e-82, Identity=55%"; Hit description:
OTHER INFORMATION: =similar to myo-inositol 1-phosphate synthase A1 [Rattus norvegicus]
FEATURE:

OTHER INFORMATION: Gene Ontology Annotation: Molecular function=inositol-3-phosphate
OTHER INFORMATION: synthase activity; Cellular component=cellular component unknown;
OTHER INFORMATION: Biological process=Myo-inositol metabolism
FEATURE:
OTHER INFORMATION: Pfam annotation: Pfam_ID=Inos-1-P_synth, Match level="Score=33.5,
OTHER INFORMATION: -value=5.8e-12, Copies=1", Pfam description=Myo-inositol-1-phosph
OTHER INFORMATION: synthase
US-60-669-175-37019

Query Match 3.3%; Score 17; DB 8; Length 280;
Best Local Similarity 100.0%; Pred. No. 7.1e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 PGLGWLVMGNGNGST 76
DB 50 PGLGWLVMGNGNGST 66

RESULT 10
US-10-990-328A-10375

Sequence 10375, Application US/10990328A
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
FILE REFERENCE: CL001495
CURRENT APPLICATION NUMBER: US/10/990,328A
CURRENT FILING DATE: 2004-11-17
NUMBER OF SEQ ID NOS: 558824
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10375
LENGTH: 558
TYPE: PRT
ORGANISM: Homo sapiens
US-10-990-328A-10375

Query Match 3.3%; Score 17; DB 6; Length 558;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 LGVWLVMGNGNGSTLT 78
DB 59 LGVWLVMGNGNGSTLT 75

RESULT 11
US-60-671-501-362

Sequence 362, Application US/60671501
GENERAL INFORMATION:
APPLICANT: Wim Van Criekinge
APPLICANT: Josef Straub
TITLE OF INVENTION: METHYLATION MARKERS FOR DIAGNOSIS AND
TITLE OF INVENTION: TREATMENT OF CANCERS
FILE REFERENCE: 006791.00001
CURRENT APPLICATION NUMBER: US/60/671,501
CURRENT FILING DATE: 2005-04-15
NUMBER OF SEQ ID NOS: 420
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 362
LENGTH: 559
TYPE: PRT
ORGANISM: Homo sapiens
US-60-671-501-362

Query Match 3.3%; Score 17; DB 8; Length 559;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 LGVWLVMGNGNGSTLT 78
DB 60 LGVWLVMGNGNGSTLT 76

RESULT 12
US-10-450-763-52092

Sequence 52092, Application US/10450763
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 52092
LENGTH: 704
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (224)..(644)
OTHER INFORMATION: Myo-inositol-1-phosphate synthase domain identified by Pfam,
OTHER INFORMATION: accession name Inos-1-P_synth, E-value=1.9e-68, Pfam score of 240.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(704)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-52092

Query Match 2.9%; Score 15; DB 6; Length 704;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 VLAHNTCEDSLAAPI 429
DB 549 VLAHNTCEDSLAAPI 563

RESULT 13
US-10-450-763-34498

Sequence 34498, Application US/10450763
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 34498
LENGTH: 237
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (1)..(237)
OTHER INFORMATION: Myo-inositol-1-phosphate synthase domain identified by Pfam,
OTHER INFORMATION: accession name Inos-1-P_synth, E-value=3.9e-07, Pfam score of -73.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(237)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-34498

Query Match 2.5%; Score 13; DB 6; Length 237;
 Best Local Similarity 100.0%; Pred. No. 8.1e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 331 SIVSYNHLGNNDG 343
 DB 170 SIVSYNHLGNNDG 182

RESULT 14
 US-10-450-763-34497
 ; Sequence 34497, Application US/10450763
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/549,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 34497
 ; LENGTH: 244
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: DOMAIN
 ; LOCATION: (1)..(244)
 ; OTHER INFORMATION: Myo-inositol-1-phosphate synthase domain identified by Pfam,
 ; OTHER INFORMATION: accession name Inos-1-P_synth, E-value=9.9e-29, Pfam score of 108
 US-10-450-763-34497

Query Match 2.5%; Score 13; DB 6; Length 244;
 Best Local Similarity 100.0%; Pred. No. 8.3e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 SIVSYNHLGNNDG 343
 DB 230 SIVSYNHLGNNDG 242

RESULT 15
 US-10-450-763-52091
 ; Sequence 52091, Application US/10450763
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/549,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 52091
 ; LENGTH: 254
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: DOMAIN
 ; LOCATION: (3)..(254)
 ; OTHER INFORMATION: Myo-inositol-1-phosphate synthase domain identified by Pfam,
 ; OTHER INFORMATION: accession name Inos-1-P_synth, E-value=4.5e-27, Pfam score of 103

US-10-450-763-52091
 Query Match 2.5%; Score 13; DB 6; Length 254;
 Best Local Similarity 100.0%; Pred. No. 8.6e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 331 SIVSYNHLGNNDG 343
 DB 240 SIVSYNHLGNNDG 252

Search completed: June 7, 2005, 17:22:47
 Job time : 43.5 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:49:11 ; Search time 6689.71 Seconds
(without alignments)
11103.899 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 atgttcacgcgaattttaa.....tgattctcgagtaacagta 1533

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hlg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_scs.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	146	9.5	1533	6	BD075269 Soybean p
2	146	9.5	1739	8	AY038802 Glycine m
3	146	9.5	1782	6	BD075266 Soybean p
4	115	7.5	1791	8	AF293970 Glycine m
5	69	4.5	2994	8	AF282263 Phaseolus
6	52	3.4	143012	8	AC125389 Medicago
7	50	3.3	62713	8	AP006418 Lotus cor
8	45	2.9	1391	8	AY005128 LOCUS
9	41	2.7	119858	8	AP006419 Actinidia
10	32	2.1	1372	8	AB090886 Aster trl
11	32	2.1	1978	8	CPINOLG
12	29	1.9	2053	8	MCU32511
13	26	1.7	1533	8	AY323824 Xerophyta
14	26	1.7	1564	8	AY096554 Arabidops
15	26	1.7	1781	6	AX054630 Sequence
16	26	1.7	1781	8	BN066307
17	26	1.7	1845	8	AY028259 Avicennia
18	26	1.7	1863	8	AY065415 Arabidops
19	26	1.7	1931	8	AB032073 Nicotiana

20	26	1.7	1950	6	E27176	E27176 Novel INPS
21	26	1.7	1959	8	AB009881	AB009881 Nicotiana
22	26	1.7	1986	8	AT333879	AT333879 Sueda ma
23	26	1.7	80088	8	AT331516	AT331516 Arabidops
24	23	1.5	35	6	BD075267	BD075267 Soybean p
25	23	1.5	1535	6	BD073472	BD073472 Regulated
26	23	1.5	1538	8	AF120146	AF120146 Triticum
27	23	1.5	1538	8	AF120147	AF120147 Triticum
28	23	1.5	1538	8	AF120148	AF120148 Triticum
29	23	1.5	1539	8	AF12340	AF12340 Porteresi
30	23	1.5	1578	6	E13612	E13612 gDNA encodi
31	23	1.5	1578	6	E17392	E17392 gDNA encodi
32	23	1.5	1665	6	BD073470	BD073470 Regulated
33	23	1.5	1665	8	AF056326	AF056326 Zea mays
34	23	1.5	1868	8	AB012107	AB012107 Oryza sat
35	23	1.5	1914	8	AK103501	AK103501 Oryza sat
36	23	1.5	1915	8	AK058750	AK058750 Oryza sat
37	23	1.5	1931	6	AR137808	AR137808 Sequence
38	23	1.5	1931	6	AR170167	AR170167 Sequence
39	23	1.5	3546	6	AR137811	AR137811 Sequence
40	23	1.5	3546	6	AR137812	AR137812 Sequence
41	23	1.5	3546	6	AR170170	AR170170 Sequence
42	23	1.5	3546	6	AR170171	AR170171 Sequence
43	23	1.5	3618	8	AF323175	AF323175 Zea mays
44	23	1.5	129658	8	AC116604	AC116604 Genomic s
45	22	1.4	682	8	AT529402	AT529402 Arabidops

ALIGNMENTS

RESULT 1	BD075269	1533 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Soybean plant producing seeds with reduced levels of raffinose				
LOCUS	BD075269				
VERSION	BD075269.1	GI:22620872			
KEYWORDS	JP 2001519665-A/4.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1533)				
AUTHORS	Hitz,W.D. and Sebastian,S.A.				
TITLE	Soybean plant producing seeds with reduced levels of raffinose				
JOURNAL	saccharides and phytic acid				
COMMENT	Patent: JP 2001519665-A 4 23-OCT-2001; EI DU PONT DE NEMOURS AND CO OS Soybean line LR33 PN JP 2001519665-A/4 PD 23-OCT-2001 PF 07-APR-1998 JP 1998543012 PR 08-APR-1997 US 08/835751 PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC Strandedness: Double; CC Topology: linear; CC Soybean plant producing seeds with reduced levels of raffinose CC and phytic acid CC Key Location/Qualifiers FH Key Location/Qualifiers FT CDS 1..1533. location/Qualifiers 1..1533 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"				
FEATURES	source				
ORIGIN	Query Match	9.5%	Score 146;	DB 6;	Length 1533;
	Best Local Similarity	100.0%	Pred. No. 6,7e-72;		
	Matches 146;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 817 ATGAAAAATGTTCTTTCAATTAATGGAAGCCCTCAGAAACACTTTTGTAACAAGGCTGATT 876
| | | | |
DB 817 ATGAAAAATGTTCTTTCAATTAATGGAAGCCCTCAGAAACACTTTTGTAACAAGGCTGATT 876
| | | | |
QY 877 GATCTTGCATCGGAGAGAACACTTTGATGTGAGATGACTTCAAGAGTGTGACAGCC 936
| | | | |
DB 877 GATCTTGCATCGGAGAGAACACTTTGATGTGAGATGACTTCAAGAGTGTGACAGCC 936
| | | | |
QY 937 AAAATGAATCTGTGTGTTGATT 962
| | | | |
DB 937 AAAATGAATCTGTGTGTTGATT 962
| | | | |
RESULT 2
LOCUS AY038802 1739 bp mRNA linear PLN 16-JUL-2001
DEFINITION Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.
ACCESSION AY038802
VERSION AY038802.1 GI:14764465
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 1739)
Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.
Biochemical and Molecular Characterization of a Mutation that
Confers a Decreased Raffinose and Phytic Acid Phenotype on
Soybean Seeds
Unpublished
2 (bases 1 to 1739)
Carlson,T.J. and Hitz,W.D.
Direct Submission
Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402,
Wilmington, DE 19880-0402, USA
JOURNAL
AUTHORS
TITLE
FEATURES
source
1. 1739
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Wye"
/db_xref="taxon:3847"
/dev_stage="developing embryo"
11. 1543
/EC_number="5.5.1.4"
/function="cyclizes glucose 6-phosphate to
L-myo-inositol-1-phosphate"
/note="mIPS-1A"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/db_xref="GI:14764466"
/translation="MEINFKVCEPNVXTETELQSYVNYETTELVHNRNGYOMIV
KPKSYEFKTIHVPLKGLVGMWGNNGSTLTGVLANBEGISMATKDKIQOANYE
GSLTQSAIRVGSFOGEBIYAPFKSLPVPNDIVPGMDISNNLDAARAVFD
IDLQKARPYMESMLPGIYDPRIANOEPRANVVKGTQBOVOOIIIDIDKAE
ATKDVXVVLTAANTERYSNLVGNDTMENTLAAVDNREBISTTAYALACMENTV
PFINSPQTFVPGILDLAIANTLIGDDPKSGQTKKSVLVDLVAGIKPISIV
YNHLGNDGNNLSAPQTRSKISKSNVDDMVNSNALYBGEHPHVIVKVPYV
GDSKRAMDEYTSIEFMGKSTIVLANTCEDSLAPILIDLVLAELSTRLEFKENE
GKFSHFHPVATILSYLTKAPLVPGTPVNVNLSKRAMLENIMRACVGLADENNNILE
YK"

QY 877 GATCTTGCATCGGAGAGAACACTTTGATGTGAGATGACTTCAAGAGTGTGACAGCC 936
| | | | |
DB 877 GATCTTGCATCGGAGAGAACACTTTGATGTGAGATGACTTCAAGAGTGTGACAGCC 946
| | | | |
QY 937 AAAATGAATCTGTGTGTTGATT 962
| | | | |
DB 947 AAAATGAATCTGTGTGTTGATT 972
| | | | |
RESULT 3
LOCUS BD075266 1782 bp DNA linear PAT 27-AUG-2002
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION BD075266
KEYWORDS BD075266.1 GI:22620869
SOURCE JP 2001519665-A/1.
ORGANISM
REFERENCE
1 (bases 1 to 1782)
Hitz,W.D. and Sebastian,S.A.
Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid
Patent: JP 2001519665-A 1 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
OS Soybean line LR13
PN JP 2001519665-A/1
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC and phytic acid
FH key Location/Qualifiers
FT CDS Location/Qualifiers
1. 1782
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 9.5%; Score 146; DB 8; Length 1739;
Best Local Similarity 100.0%; Pred. No. 6.7e-72;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGAAAAATGTTCTTTCAATTAATGGAAGCCCTCAGAAACACTTTTGTAACAAGGCTGATT 876
| | | | |
DB 827 ATGAAAAATGTTCTTTCAATTAATGGAAGCCCTCAGAAACACTTTTGTAACAAGGCTGATT 886
| | | | |

RESULT 4
LOCUS AF293970 1791 bp mRNA linear PLN 03-MAY-2001
DEFINITION Glycine max myo-inositol-3-phosphate synthase (MIPS) mRNA, complete
cds.
ACCESSION AF293970
VERSION AF293970.1 GI:13936690

rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.

1 (bases 1 to 2994)

Johnson, M.D., Lackey, R.H. and Pope, P.M.
Subcellular localization of 1L-myo-inositol-1-phosphate synthase
Unpublished

2 (bases 1 to 2994)

Pope, P.M. and Johnson, M.D.
Direct Submission
Submitted (24-JUN-2000) Biological Sciences, The University of
Alabama, P.O. Box 870344, Tuscaloosa, AL 35487, USA
location/Qualifiers

source

1. 2994

/organism="Phaseolus vulgaris"

/mol_type="genomic DNA"

/db_xref="taxon:3885"

join(-582, 769, 886, .954, 1047, .1182, 1335, .1582, 1662, .1888, 1965, .2080, 2243, .2419, 2503, .2691, 2795, .>2860)

/product="1L-myo-inositol-1-phosphate synthase"

join(582, .769, 886, .954, 1047, .1182, 1335, .1582, 1662, .1888, 1965, .2080, 2243, .2419, 2503, .2691, 2795, .>2860)

/codon_start=1

/product="1L-myo-inositol-1-phosphate synthase"

/protein_id="AAK69514.1"

/db_xref="GI:14582467"

/translation="MFIENFKVESPNVKYSETEIQSVNYETTELVHENGAYQWII
KPSVAFKPTNTHVPRKLGVMVGWGNNGSITLGGVIANREBISVATKDKIQANF
GSLTQSAIRPVESGGEIYAPRKSIPDMVNPDIYVGGDDISNNMLADMGAKVFD
IDLOKRLPRVMSMVPDLPGIYDPDFIAANDEEANNVYKGTKKRQVOQVKKDKPEFA
ATYDQVIVLMTANTERTERSNLVYGLADPTENLLAALDRNFAELSPSTPAIACTEVY
PFINSGPDDVTVFVGLIDFAIQKNCILGDDDFKSGQTKMKSLVDFELVGAIGKPTSYIS
YNLGNNDGNNLSAPQTFPSKELISKSNVVDVWNSNALIYEPEDHPDHVVAVIKVPIY
AASKRAMDEYTSIEIFMGGRKTVITVANTCEBDSLAAPIIDLVLLAEISTRIQFAEKGE
GKHSHPVATILSYLTAKPLV"

ORIGIN

Query Match 4.5%; Score 69; DB 8; Length 2994;
Best local Similarity 100.0%; Pred. No. 2e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1093 GATATGTCACAGCAATGCATCCCTATAGACCGTGTAACATCCCGACCATGTGTT 1152
|||||
Db 2351 GATATGTCACAGCAATGCATCCCTATAGACCGTGTAACATCCCGACCATGTGTT 2410
|||||

QY 1153 GTTATTAG 1161
|||||

Db 2411 GTTATTAG 2419

RESULT 6
AC125389/c 143012 bp DNA linear PLN 15-JUN-2004

LOCUS
DEFINITION
AC125389 Medicago truncatula clone mch2-12a18, complete sequence.
AC125389
VERSION
KEYWORDS
HTG. GI:48717535
ORGANISM
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 143012)
Shauli, S., Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Medicago truncatula BAC Clone mch2-12a18
Unpublished
2 (bases 1 to 143012)
Shauli, S., Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (25-JUN-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

REFERENCE
AUTHORS
TITLE
JOURNAL
OK 73019, USA
3 (bases 1 to 143012)
Shall, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (20-MAY-2004) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 143012)
Shall, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (15-JUN-2004) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Jun 15, 2004 this sequence version replaced gi:47524030.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code: UOKNOR

FEATURES
source
1..143012
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mtn2-12a18"
/clone_lib="Medicago truncatula BAC library H2"
ORIGIN
Query Match 3.4%; Score 52; DB 8; Length 143012;
Best Local Similarity 100.0%; Pred. No. 1.2e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1072 TCCAGAGCAAGCTGTGACATATGTCACACGCAATGCTCTCATG 1123
|||||
Db 14603 TCCAGAGCAAGCTGTGACATATGTCACACGCAATGCTCTCATG 14552
RESULT 7
AP006418 62713 bp DNA linear PLN 22-JUL-2003
LOCUS
DEFINITION
Lotus corniculatus var. japonicus genomic DNA, chromosome 6,
clone:ljt3A02, TM0306, complete sequence.
AP006418
AP006418.1 GI:31581049
HTG.
Lotus corniculatus var. japonicus (Lotus japonicus)
SOURCE
Lotus corniculatus var. japonicus
ORGANISM
Lotus corniculatus var. japonicus
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;
Lotus.
REFERENCE
AUTHORS
TITLE
JOURNAL
Sato, S., Kato, T., Sato, S., Nakamura, Y., Kaneko, T. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. IV. Sequence
Features and Mapping of seventy-three TAC clones which cover the
7.5 Mb Regions of the Genome
DNA Res. (2003) In press
2 (bases 1 to 62713)
Sato, S.
Direct Submission
Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research: 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex.2337),
Fax: 81-438-52-3934)
FEATURES
source
1..62713
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"

REFERENCE
AUTHORS
TITLE
JOURNAL
OK 73019, USA
3 (bases 1 to 143012)
Shall, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (20-MAY-2004) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 143012)
Shall, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (15-JUN-2004) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Jun 15, 2004 this sequence version replaced gi:47524030.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code: UOKNOR

FEATURES
source
1..143012
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mtn2-12a18"
/clone_lib="Medicago truncatula BAC library H2"
ORIGIN
Query Match 3.4%; Score 52; DB 8; Length 143012;
Best Local Similarity 100.0%; Pred. No. 1.2e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1072 TCCAGAGCAAGCTGTGACATATGTCACACGCAATGCTCTCATG 1123
|||||
Db 14603 TCCAGAGCAAGCTGTGACATATGTCACACGCAATGCTCTCATG 14552
RESULT 7
AP006418 62713 bp DNA linear PLN 22-JUL-2003
LOCUS
DEFINITION
Lotus corniculatus var. japonicus genomic DNA, chromosome 6,
clone:ljt3A02, TM0306, complete sequence.
AP006418
AP006418.1 GI:31581049
HTG.
Lotus corniculatus var. japonicus (Lotus japonicus)
SOURCE
Lotus corniculatus var. japonicus
ORGANISM
Lotus corniculatus var. japonicus
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;
Lotus.
REFERENCE
AUTHORS
TITLE
JOURNAL
Sato, S., Kato, T., Sato, S., Nakamura, Y., Kaneko, T. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. IV. Sequence
Features and Mapping of seventy-three TAC clones which cover the
7.5 Mb Regions of the Genome
DNA Res. (2003) In press
2 (bases 1 to 62713)
Sato, S.
Direct Submission
Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research: 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex.2337),
Fax: 81-438-52-3934)
FEATURES
source
1..62713
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"

REFERENCE
AUTHORS
TITLE
JOURNAL
OK 73019, USA
3 (bases 1 to 143012)
Shall, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (20-MAY-2004) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 143012)
Shall, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (15-JUN-2004) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Jun 15, 2004 this sequence version replaced gi:47524030.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code: UOKNOR

FEATURES
source
1..1391
/organism="Actinidia arguta"
/mol_type="mRNA"
/db_xref="taxon:64478"
/tissue_type="young mature leaf"
1..1095
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAP97409.1"
/translation="MNLADAVARARVFDIDLOQLRPYMSWPLPGIYEPPDPAANO
GSRANNVYKSGKKEQVQIIDIREFEKKNKRVVVLMTANTERYSNVYGLNDTME
NLPSLEKNSEIISPTLYAACVLENI PFINGSPQNFVGLIDLAIRNCLGDD
FKSQTKMSKLVLPVAGAGIKPTISYVNLGNDQNLISA POTFSKESKSNVVD
DMVSSNATVYRGRGPHVVIKYVPVAVGSKRAMDYMSEIFMGKNITVLTNCE
SLAAPITIDVLAELSTRIOQLKADDEGKSHSPVATILSYLTAKVLPVPGIPVNV
ALSKORAMLENTLACIGLSPENNMLEYK"
ORIGIN
Query Match 2.9%; Score 45; DB 8; Length 1391;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 900 TTTGATTGTCGAGATGACTTCAAGAGTGTGACGCAAAATGAA 944
|||||
Db 462 TTTGATTGTCGAGATGACTTCAAGAGTGTGACGCAAAATGAA 506
RESULT 9
AP006419 119858 bp DNA linear PLN 22-JUL-2003
LOCUS
DEFINITION
Lotus corniculatus var. japonicus genomic DNA, chromosome 4,
clone:ljt3018, TM0307, complete sequence.
AP006419
AP006419.1 GI:31581050

KEYWORDS HTG.
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

REFERENCE 1
AUTHORS Asamizu, E., Kato, T., Sato, S., Nakamura, Y., Kaneko, T. and Tabata, S.
TITLE Structural Analysis of a Lotus japonicus Genome. IV. Sequence Features and Mapping of seventy-three TAC clones which cover the 7.5 Mb Regions of the Genome
JOURNAL DNA Res. (2003) In press
REFERENCE 2 (bases 1 to 119858)
AUTHORS Sato, S.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp, URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337), Fax: 81-438-52-3934)

FEATURES
source
1. Location/Qualifiers
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="4"
/clone="LJT43018"
/clone_11b="LJT 1library"
/note="TAC clone: TW0307-synonym: Lotus japonicus"

ORIGIN
Query Match 2.7%; Score 41; DB 8; Length 119858;
Best Local Similarity 100.0%; Pred. No. 2,8e-11;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 AGGCTTACATGGAATCCATGTTCCATCCCGGAAATCTA 542
Db 47386 AGGCTTACATGGAATCCATGTTCCATCCCGGAAATCTA 47346

RESULT 10
LOCUS AB090886 1372 bp mRNA linear PLN 04-MAR-2003
DEFINITION Aster tripolium mRNA for myo-inositol-1-phosphate synthase, partial cds.
ACCESSION AB090886
VERSION AB090886.1 GI:28804514
KEYWORDS Aster tripolium
SOURCE Aster tripolium
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Campanulids; Asterales; Asteraceae; Asteroideae; Astereae; Aster.

REFERENCE 1
AUTHORS Takeda, M., Uno, Y., Kanuchi, M. and Inagaki, N.
TITLE Analyze of nine cDNAs for salt-inducible genes in the halophyte sea aster
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1372)
AUTHORS Takeda, M., Uno, Y., Kanuchi, M. and Inagaki, N.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2002) Migiwa Takeda, The Graduate School of Science and Technology, Kobe University, Division of Environmental Science, Rokkondaiyuu 1-1, Kobe-shi Nada-ku 657-8501, Japan (E-mail: 991d860n@01.kobe-u.ac.jp, Tel: 81-78-803-5832, Fax: 81-78-803-5832)

FEATURES
source
1. Location/Qualifiers
/organism="Aster tripolium"
/mol_type="mRNA"

CDS
/db_xref="taxon:74787"
<1..1230
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="BAC57963.1"
/db_xref="GI:28804515"

ORIGIN
Query Match 2.1%; Score 32; DB 8; Length 1372;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 AGGCTTACATGGAATCCATGTTCCATCC 533
Db 202 AGGCTTACATGGAATCCATGTTCCATCC 233

RESULT 11
LOCUS CPINOIG 1978 bp DNA linear PLN 26-JUN-1995
DEFINITION C. paradiisi (Macf) INO1 gene.
ACCESSION 232632
VERSION 232632.1 GI:602564
KEYWORDS INO1 gene.
SOURCE Citrus x paradiisi
ORGANISM Citrus x paradiisi
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Sapindales; Rutaceae; Citrus.

REFERENCE 1 (bases 1 to 1978)
AUTHORS Holland, D.
TITLE Direct Submission
JOURNAL Submitted (08-APR-1994) Doron Holland, Fruit Tree Breeding and Genetics, Volcani Research, Organisation Center, Bet Dagan, 50250, Israel

FEATURES
source
1. Location/Qualifiers
/organism="Citrus x paradiisi"
/mol_type="genomic DNA"
/db_xref="taxon:37656"
/clone="c38"
/class_type="fully expanded leaves"
/dev_stage="juvenile plants"
/gene="INO1"
/gene_11b="INO1"
/gene="INO1"
/codon_start=1
/protein_id="CAA83565.1"
/db_xref="GI:602565"
/db_xref="GOA:P42802"
/db_xref="UniProt/Swiss-Prot:P42802"
/translation="PF1ENPVSPPVNTDHEHSVVDYETTELVHNRNTQYMI
RKRTYREKTDVHVPRKGVMLVGGNNGSLTGTGIVANRGTICWATKDVQANYP
GSLTQSAIRGVSNGEIEIVAFKSLILMVNPDDIVFGWMDISDNLADNANRAVFD
IDLQKRLRPYSESWPLPGIYDPPRIANQSRANNVIKGTKEQMLQIIXDIREFKE
KKIVDRVVLWNTANERSNVVIGNDIVESLSLADIKKEAISSTLYALACVLENI
PPIINSPPVLPGLIDAIRNCLIGDDDFSGQTKKSVLVDFLVAGAIKPTISV
VNHGNNQNMNLASQTRSKSISNVVDMVSNVPMGLVNTPRPWIKYVPPVAVI
ERANDYTSSEIFMGSKTIVLHNTEDSLAIPALIIDLVLAELSTRYQLKRGEGK
HSFHPVATILSLTAPLVPPTPVNALSKORALENIIIRACVGLAENNNMILEYK"

ORIGIN
Query Match 2.1%; Score 32; DB 8; Length 1978;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 ATGCTTGTTGGGTGGTGAACAACGCGCTC 224
|||||
Db 380 ATGCTTGTTGGGTGGGTGAACAACGCGCTC 411

RESULT 12
MCU32511 2053 bp mRNA linear PLN 11-JUL-1996
LOCUS Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase
DEFINITION
ACCESSION U32511
VERSION U32511.1 GI:975887
KEYWORDS
SOURCE
ORGANISM Mesembryanthemum crystallinum (common iceplant)
REFERENCE
AUTHORS Mesembryanthemum crystallinum
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Aizoaceae; Mesembryanthemum.
TITLE 1 (bases 1 to 2053)
AUTHORS Ishihara, M., Majumder, A. L., Bornhouser, A., Michalowski, C. B.,
Jensen, R. G., and Bohner, H. J.
COORDINATE TRANSCRIPTIONAL INDUCTION OF MYO-INOSITOL METABOLISM
DURING ENVIRONMENTAL STRESS
PLANT J. 9 (4), 537-548 (1996)
JOURNAL 96208959
MEDLINE 8624516
PUBMED
REFERENCE 2 (bases 1 to 2053)
AUTHORS Ishihara, M., Majumder, A. L., Bornhouser, A., Michalowski, C. B.,
Jensen, R. G., and Bohner, H. J.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1995) Hans J. Bohner, Biochemistry, University
of Arizona, Biosciences West 516, Tucson, AZ 85721, USA
FEATURES
source
1..2053
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
182..1720
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AA03687.1"
/db_xref="GI:975888"
/translation="MFISFVYSPNVKXTENIEISVYNYDTTELVHNRKDAQGYOM
IVKPTVOYHFKTDRVPLKGLVGMVGMKNGNSTLGLGVINREGISWATDKIOAN
YRGSITQASSTIRVSPNGEETIAPFKSLPMPNPDVYFGNDIDNMLADMTARV
FDIDLOKQRPYMEHWPUGIYDPDFIANQSRANVTKTKKEQVERVKDIREF
KEKNVYDRVYVMTGENTERYSNVVGLDNTMENLASLEKKESEISPSSTVALCIEE
NIPFINGSPONTFVPGQLIDLAIKNSLIGQDPFKSGQTKMSVLVDFVGAIGKPTSV
VYNHIGNDGNMLSAPOTFRSEKISKSVVDVWASNGIILEPEHDPHVAVIKYVP
VYGSKRAMEYTESIFMGCGTNTIVMHTCESLIAAPILIDVLVLAELSTRIOKAE
BEDKRSHPVATITISYLTKAPLVPPGTPVVALAKORAMLENITRACVGLAPENNMI
LEYK"

ORIGIN
Query Match 1.9%; Score 29; DB 8; Length 2053;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1183 AGCAAGAGCCATGCATGATGATGACTTC 1211
|||||
Db 1370 AGCAAGAGCCATGCATGATGATGACTTC 1398

RESULT 13
AY333824 1533 bp mRNA linear PLN 15-JUL-2003
LOCUS Xerophyta viscosa myo-inositol-1-phosphate synthase INO1 mRNA,
DEFINITION complete cds.
ACCESSION AY333824
VERSION AY333824.1 GI:32492881
KEYWORDS

SOURCE Xerophyta viscosa
ORGANISM Xerophyta viscosa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Velloziaceae; Xerophyta.
REFERENCE 1 (bases 1 to 1533)
AUTHORS Majee, M., Majumder, A. N. L., and Munder, S. G.
TITLE Molecular characterization of Xvino1, a myo-inositol-1-phosphate
synthase from the resurrection plant Xerophyta viscosa
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1533)
AUTHORS Majee, M., Majumder, A. N. L., and Munder, S. G.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-2003) Molecular and Cell Biology, University of
Cape Town, Private Bag, Rondebosch, Western Cape 7701, South Africa
FEATURES
source
1..1533
/organism="Xerophyta viscosa"
/mol_type="mRNA"
/db_xref="taxon:90708"
1..1533
/note="Xvino1"
/codon_start=1
/product="myo-inositol-1-phosphate synthase INO1"
/protein_id="AB8553.1"
/db_xref="GI:32492882"
/translation="MFIESPVRVSPHYRGAAEIESDYQYDTTELVHSHDASRWV
RPSVRVNFRTTTPVPLKGLVGMVGMKNGNSTLGLGVINREGISWATDKIOQANY
GLTQASTIRVSPNGEETIAPFKSLPMPNPDVYFGNDIDNMLADMTARVLD
IDLOKQRPVMEHWPUGIYDPDFIANQSRANVTKTKKEQVERVKDIREF
KSKYDRVYVMTGENTERYSNVVGLDNTMENLASVDRKKESEISPSSTVALCMEGI
PFINGSPONTFVPGQLIDLAIKNSLIGQDPFKSGQTKMSVLVDFVGAIGKPTSV
VYNHIGNDGNMLSAPOTFRSEKISKSVVDVWASNGIILEPEHDPHVAVIKYVPV
GDSKRAMEYTESIFMGCGTNTIVMHTCESLIAAPILIDVLVLAELSTRIOKAE
EKPSRHPVATITISYLTKAPLVPPGTPVVALAKORAMLENITRACVGLAPENNMI
LEYK"

ORIGIN
Query Match 1.7%; Score 26; DB 8; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1105 AGCAATGCCATCCTCTATGACCTGG 1130
|||||
Db 1105 AGCAATGCCATCCTCTATGACCTGG 1130

RESULT 14
AY096554 1564 bp mRNA linear PLN 18-SEP-2002
LOCUS Arabidopsis thaliana putative myo-inositol-1-phosphate synthase
DEFINITION (AF591017) mRNA, complete cds.
ACCESSION AY096554
VERSION AY096554.1 GI:20465488
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1564)
AUTHORS Yamada, K., Banh, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M.,
Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L.,
Tang, C., Toriumi, M., Wu, H. C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamliya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Meyers, M. C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J.,
Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A.,
Shinozaki, K., Davis, R. W., Ecker, J. R., and Theologis, A.
TITLE Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1564)
AUTHORS Yamada, K., Banh, J., Chan, M. M., Chang, C. H., Chang, E., Dale, J. M.,
Deng, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L.,

Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission

Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : RIKEN Arabidopsis Full-length cDNA1) : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PEGC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Banb, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PEGC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PEGC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES
source
Location/Qualifiers

1..1564
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U18590"
/ecotype="Columbia"
/note="this clone is in pUNI 51."
1..1564
/gene="At5g10170"
1..1533
/gene="At5g10170"
/codon_start=1
/evidence=experimental
/product="putative myo-inositol-1-phosphate synthase"
/protein_id="AA020204.1"
/db_xref="GI:20465489"
/translation="MTESPKVESPVKXTENHSVVDYOTTELVHKNKGAFOKTVKRTVYEFKTDTHVFKLGMVYKGGNGSTITRAGVIANREGISWATKEKYQOANYF GSLTQASIRVSPNGEETIYAPPKSLIPVNPEELVEGWDISDMNLADAMARAVLD IDLOKOMRPFMEHWPLPGIFDPDFTAAQGSANVIGTKKQOLEVYIKDIRKEKE KKRVKVVLTANTERYSNVVGLNDITENLMSLIDKDEAESPTSLVIAICVLENY PFINGSPONTFWGLIELAIKKNCLIGDDPKSGOTKMSVLVDVFGAGIKPTSIYS VNHIGNNDGMNL SAPOTFSKISNSNVDDWVSGNGLIYEPGEHDPHVVIKYVPCV GSKRAMDEYTSIEIPFGKRTTYMARTCDLSLAAPITLIDVLIALELTTRIQPMSENE GKFSHFHPVATLISLTKAPLVPGTGVNALSKQAPMLNVLRACVGLAPENNMLE YK"

3'UTR
ORIGIN
Query Match 1.7%; Score 26; DB 8; Length 1564;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1534..1564
/gene="At5g10170"

QY 913 GATGACTCAAGAGTGTCAAGACCA 938
DB 913 GATGACTCAAGAGTGTCAAGACCA 938
RESULT 15
AX054630

LOCUS AX054630 1781 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 1 from Patent WO0073473.
ACCESSION AX054630
VERSION AX054630.1 GI:12228190
KEYWORDS
SOURCE
ORGANISM

Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
1
Georges, F., Hussain, A.A. and Keller, W.A.
Method for reducing phytate in canola meal using genetic manipulation involving myo-inositol 1-phosphate synthase gene Patent: WO 0073473-A 1 07-DEC-2000;
NATIONAL RESEARCH COUNCIL CANADA (CA)

FEATURES
source
Location/Qualifiers
1..1781
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"

ORIGIN
Query Match 1.7%; Score 26; DB 6; Length 1781;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 907 GGTTGAGATGACTTCAAGAGTGTCA 932
DB 963 GGTTGAGATGACTTCAAGAGTGTCA 988

Search completed: June 8, 2005, 11:26:38
Job time : 6690.71 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:47:21 ; Search time 842.149 Seconds
(without alignments)
10775.958 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgttcacgcgagaattctaa.....tgattctcgagracaaagta 1533

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1533	100.0	1533	12	ADQ14500	Adq14500 Mutant so
2	1533	100.0	1533	13	AD882001	Ad882001 Soybean m
3	1482	96.7	1533	12	ADQ14504	Adq14504 Wild type
4	1482	96.7	1533	12	ADQ14502	Adq14502 Mutant so
5	1482	96.7	1533	13	AD882005	Ad882005 Soybean m
6	1482	96.7	1533	13	AD882003	Ad882003 Soybean m
7	146	9.5	1533	2	AAV62443	AAV62443 Soybean m
8	146	9.5	1533	12	ADQ14498	Adq14498 Mutant so
9	146	9.5	1533	12	ADQ14494	Adq14494 Mutant so
10	146	9.5	1533	13	AD881999	Ad881999 Soybean m
11	146	9.5	1533	13	AD881993	Ad881993 Soybean m
12	146	9.5	1533	13	AD881997	Ad881997 Soybean m
13	146	9.5	1760	12	ADQ14490	Adq14490 Wild type
14	146	9.5	1782	2	AAV62440	AAV62440 Soybean w
15	26	1.7	377	2	ADP93370	Adp93370 Cotton ex
16	26	1.7	532	13	ACN61638	Acn61638 Cotton gy
17	26	1.7	533	13	ACN62170	Acn62170 Cotton gy
18	26	1.7	543	13	ACN61870	Acn61870 Cotton gy
19	26	1.7	566	13	ACN57874	Acn57874 Cotton gy
20	26	1.7	618	13	ACN62990	Acn62990 Cotton ca

21	26	1.7	1781	4	AC87643	Aac87643 Brassica
22	26	1.7	1950	2	AAV90402	Aav90402 Nicotiana
23	23	1.5	35	2	AAV62441	Aav62441 Soybean m
24	23	1.5	35	12	ADQ14492	Adq14492 Soybean m
25	23	1.5	345	12	ADP94497	Adp94497 Cotton ex
26	23	1.5	452	12	ADP94561	Adp94561 Cotton ex
27	23	1.5	557	13	ACN58059	Acn58059 Cotton gy
28	23	1.5	601	13	ACN60643	Acn60643 Cotton gy
29	23	1.5	605	13	ACN60421	Acn60421 Cotton gy
30	23	1.5	1536	12	ADP43918	Adp43918 P. coarct
31	23	1.5	1578	2	AAV91656	Aav91656 Inositol-
32	23	1.5	1578	2	AAV65400	Aav65400 Inducible
33	23	1.5	1665	2	AAV09006	Aav09006 Maize myo
34	23	1.5	1931	2	AAV24407	Aav24407 Maize myo
35	23	1.5	3546	2	AAV24410	Aav24410 Maize myo
36	23	1.5	3546	2	AAV24411	Aav24411 Maize myo
37	21	1.4	35	13	AD881995	Ad881995 PCR prime
38	21	1.4	39	12	ADQ14493	Adq14493 Soybean m
39	21	1.4	39	12	ADQ14493	Adq14493 Soybean m
40	20	1.3	291	12	ADP93783	Adp93783 Cotton ex
41	20	1.3	388	13	ACN52468	Acn52468 Cotton an
42	20	1.3	389	13	ACN51365	Acn51365 Cotton an
43	20	1.3	449	12	ADP91072	Adp91072 Cotton ex
44	20	1.3	506	13	ACN57963	Acn57963 Cotton gy
45	20	1.3	517	13	ACN63079	Acn63079 Cotton ca

ALIGNMENTS

RESULT 1	ADQ14500	ADQ14500 standard; cDNA, 1533 BP.
ID	ADQ14500	
XX	ADQ14500	
AC	ADQ14500	
XX		
DT	23-SEP-2004	(first entry)
XX		
DE	Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.	
XX		
KW	Soybean; myo-inositol 1-phosphate synthase; gene; ss;	
KW	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;	
XX	inorganic phosphate; mutant.	
XX		
OS	Glycine max.	
XX		
XX	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1533
FT		/*tag= a
FT		/product= "Mutant soybean myo-inositol 1-phosphate
FT		synthase #3"
XX		
PN	US2004128713-A1.	
XX		
PD	01-JUL-2004.	
XX		
XX		
PF	21-NOV-2003; 2003US-00718952.	
XX		
PR	08-APR-1997; 97US-00835751.	
PR	07-APR-1998; 98WO-US006822.	
PR	26-APR-1999; 99US-00299315.	
PR	11-MAR-2002; 2002US-00025003.	
XX		
PA	(HITZ/) HITZ W D.	
PA	(SEBA/) SEBASTIAN S A.	
PA	(GRAC/) GRACE D J.	
PA	(STRE/) STREIT L G.	
XX		
XX	Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
PI	WPI; 2004-533135/51.	
DR	P-Psdb; ADQ14501.	

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Claim 10; SEQ ID NO 11; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 100.0%; Score 1533; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCATCGAGAAATTTAAGTAGAGAGTCTTAATGGAATACCGAGCTGAGATT 60
DB 1 ATGTCATCGAGAAATTTAAGTAGAGAGTCTTAATGGAATACCGAGCTGAGATT 60
QY 61 CAGTCGCTGTACACTACGAAACCGAAGCTGTTTCAAGAGAGAGATGGACCTAT 120
DB 61 CAGTCGCTGTACACTACGAAACCGAAGCTGTTTCAAGAGAGAGATGGACCTAT 120
QY 121 CAGTGGATTTGCAAAACCCAAATCCGTCACATCAATTTAAACCAACCCATGTTCA 180
DB 121 CAGTGGATTTGCAAAACCCAAATCCGTCACATCAATTTAAACCAACCCATGTTCA 180
QY 181 AAATTGGGGGTGATGCTTGGGGTGGGGTGAACAACGGCTCAACCTCCGCTGAT 240
DB 181 AAATTGGGGGTGATGCTTGGGGTGGGGTGAACAACGGCTCAACCTCCGCTGAT 240
QY 241 GTTATTGGCTAACAGAGAGACATTTTCATGGCTACAAAGACAAATTCACCAAT 300
DB 241 GTTATTGGCTAACAGAGAGACATTTTCATGGCTACAAAGACAAATTCACCAAT 300
QY 301 TACTTTGGCTCCCTCAACCAAGCTCCTATTCAGATTTGATCTTCCAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCAAGCTCCTATTCAGATTTGATCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCATTCAGAGAGTCTGCTTCAATGTTAATCCAGAGACATTTGTTGGG 420
DB 361 ATCTATGCCCATTCAGAGAGTCTGCTTCAATGTTAATCCAGAGACATTTGTTGGG 420
QY 421 GGATGGGATATACAGCAACATGAACCTGCTGATGCAATGCGCAAGGCAAGGTTTGA 480
DB 421 GGATGGGATATACAGCAACATGAACCTGCTGATGCAATGCGCAAGGCAAGGTTTGA 480
QY 481 ATCGATTTGACAGACAGCTTGAAGGCTTATCATGGAATTCATCTCCCGGAATC 540
DB 481 ATCGATTTGACAGACAGCTTGAAGGCTTATCATGGAATTCATCTCCCGGAATC 540
QY 541 TACGACCCGGATTTTATGCTGCTCAACCAAGAGGCGTCCCAACACGTGTTAAGGCG 600
DB 541 TACGACCCGGATTTTATGCTGCTCAACCAAGAGGCGTCCCAACACGTGTTAAGGCG 600

QY 601 ACAAAGCAGAGCAAGTTACGAAATCATCAAGACATCAAGGCGTTTAAAGAACCCACC 660
DB 601 ACAAAGCAGAGCAAGTTACGAAATCATCAAGACATCAAGGCGTTTAAAGAACCCACC 660
QY 661 AAAGTGGCAAGGTGTTGCTGCTGATGCTGCAACAGAGAGATTAAGAAATTTGGTT 720
DB 661 AAAGTGGCAAGGTGTTGCTGCTGATGCTGCAACAGAGAGATTAAGAAATTTGGTT 720
QY 721 GTAGGCTTTATGACACCATGAGAAATCTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTTATGACACCATGAGAAATCTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTGATGATGCTGCTGATGATGATGATGATGATGATGATGAT 840
DB 781 ATTTCTCTTCCACCTGATGATGCTGCTGATGATGATGATGATGATGATGATGAT 840
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGGGCTGATGATGATGATGATGATGATGAT 900
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGGGCTGATGATGATGATGATGATGATGAT 900
QY 901 TTGATTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 TTGATTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 TTTCTTGGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 TTTCTTGGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 AAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 AAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GACCATGTTGTTTATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GACCATGTTGTTTATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 GAGTACATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 GAGTACATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 GAGATTCCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 GAGATTCCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 ACTAGAAATCAAGTTTAAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
DB 1321 ACTAGAAATCAAGTTTAAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
QY 1381 ACCATTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 ACCATTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 GCATTTGCAAGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 GCATTTGCAAGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 CCAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1533
DB 1501 CCAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1533

RESULT 2
ADS82001
ID ADS82001 standard; cDNA: 1533 BP.
XX
XX ADS82001;
AC
XX
XX
DT 18-NOV-2004 (first entry)
XX

DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
 XX
 XX Soybean, plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
 OS Glycine max; line 29010CP01.
 XX
 XX Synthetic.
 FT Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 FT mutation replace(260,G)
 FT /tag= b
 PN US2003074685-A1.
 PD 17-APR-2003.
 PD 11-MAR-2002; 2002US-00025003.
 PF 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 XX (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PI Hitz WD, Sebastian SA;
 DR WPI; 2004-639957/62.
 DR P-PSDB; ADS82002.
 XX
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 PS Claim 8; SEQ ID NO 11; 34pp; English.
 XX
 XX The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g), a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with a
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence encodes a mutant myo-
 CC inositol 1-phosphate synthase.
 XX
 XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1533; DB 13; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0;

	Matches 1533;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	ATGTCATCGAATAATTTTAAGTAGAGTCTTAATGTGAATACCCGAGCTGAGATT							60
DB	1	ATGTCATCGAATAATTTTAAGTAGAGTCTTAATGTGAATACCCGAGCTGAGATT							60
QY	61	CAGTCGCTGTACACATGAGAAACACCCGACTTGTTCAGAGAACGGAATGGACCTAT							120
DB	61	CAGTCGCTGTACACATGAGAAACACCCGACTTGTTCAGAGAACGGAATGGACCTAT							120
QY	121	CAGTGATTTGTCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCAACCTATCCA							180
DB	121	CAGTGATTTGTCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCAACCTATCCA							180
QY	181	AAATTGGGGGTGATGCTTGTGGGTTGGGGTGAACCAACGCTCTACCCCTCACCGGTGT							240
DB	181	AAATTGGGGGTGATGCTTGTGGGTTGGGGTGAACCAACGCTCTACCCCTCACCGGTGT							240
QY	241	GTTATTTCTTAACAGAGAGACATTTCAATGGGCTTACAAAGACAAAGATTCAACAGCCAT							300
DB	241	GTTATTTCTTAACAGAGAGACATTTCAATGGGCTTACAAAGACAAAGATTCAACAGCCAT							300
QY	301	TACTTTGGCTCCCTCAACCAACCTCAGCTATTGAGTTGATCTTCCAGGAGAGAGAA							360
DB	301	TACTTTGGCTCCCTCAACCAACCTCAGCTATTGAGTTGATCTTCCAGGAGAGAGAA							360
QY	361	ATCTATGCCCATTCAGAGAGTCTGCTTCCATGTGTTAATCTTGACGACATTGTGTTGGG							420
DB	361	ATCTATGCCCATTCAGAGAGTCTGCTTCCATGTGTTAATCTTGACGACATTGTGTTGGG							420
QY	421	GGATGGGANTATGACAACATGAACTGGCTGATGGCAATGGCCAGGCAAAAGGTGTTGAC							480
DB	421	GGATGGGANTATGACAACATGAACTGGCTGATGGCAATGGCCAGGCAAAAGGTGTTGAC							480
QY	481	ATCGATTTGACAGAACAGTTGAGGCTTACATGATCATGATGTTCCACTCCCGGAATC							540
DB	481	ATCGATTTGACAGAACAGTTGAGGCTTACATGATCATGATGTTCCACTCCCGGAATC							540
QY	541	TACGACCCGGATTTTCAATGCTGTCACCAACAGAGAGCGTGCCACACGCTATTAAGGCG							600
DB	541	TACGACCCGGATTTTCAATGCTGTCACCAACAGAGAGCGTGCCACACGCTATTAAGGCG							600
QY	601	AAAAAGCAAGACCAAGTTCAGCAATCATCAAGATCAAAAGCCGTTTAAAGAAAGCAAC							660
DB	601	AAAAAGCAAGACCAAGTTCAGCAATCATCAAGATCAAAAGCCGTTTAAAGAAAGCAAC							660
QY	661	AAAGTGACAAAGGTGTTGTCCTGTGACCTGCAACACAGAGAGGTATGCAATTTGGTT							720
DB	661	AAAGTGACAAAGGTGTTGTCCTGTGACCTGCAACACAGAGAGGTATGCAATTTGGTT							720
QY	721	GTAGGCTTTAATGACACATGAGAAATCTTTGCTGTGTGACAGAAATGAGGCTGAG							780
DB	721	GTAGGCTTTAATGACACATGAGAAATCTTTGCTGTGTGACAGAAATGAGGCTGAG							780
QY	781	ATTTTCTCTTCCACTTGTATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATAT							840
DB	781	ATTTTCTCTTCCACTTGTATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATAT							840
QY	841	GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACCT							900
DB	841	GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACCT							900
QY	901	TTGATTTGGTAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGAT							960
DB	901	TTGATTTGGTAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGAT							960
QY	961	TTTCTTTGGGGGCTGTGTATCAAGCAACATCTAATAGTTAAGTTAACCATCTGTGAAAC							1020
DB	961	TTTCTTTGGGGGCTGTGTATCAAGCAACATCTAATAGTTAAGTTAACCATCTGTGAAAC							1020
QY	1021	AATGATGATGATATCTCTCGGCTCCAAACCTTCCGCTCCAAAGAAATCTCCAAAGAC							1080
DB	1021	AATGATGATGATATCTCTCGGCTCCAAACCTTCCGCTCCAAAGAAATCTCCAAAGAC							1080

QY 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCCTTATGAGCCTGTGAAACATCC 1140
DB 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCCTTATGAGCCTGTGAAACATCC 1140
QY 1141 GACCATGTTGTTGTTATTAATGATATGCTTTCGTAAGGAGATGCAAGAGACCATGAT 1200
DB 1141 GACCATGTTGTTGTTATTAATGATATGCTTTCGTAAGGAGATGCAAGAGACCATGAT 1200
QY 1201 GAGTACACTTCGAGATATTCATGAGTGGGAAAGAACACCATGTTTGGACAACATATG 1260
DB 1201 GAGTACACTTCGAGATATTCATGAGTGGGAAAGAACACCATGTTTGGACAACATATG 1260
QY 1261 GAGGATTCCTTTTGAAGCTCTTATTAATTCCTTGAAGCTTCTTCTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTGAAGCTCTTATTAATTCCTTGAAGCTTCTTCTGAGCTGAGC 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATTCACCCAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATTCACCCAGTTGCT 1380
QY 1381 ACCATTCTCAGTATCTGACCAAGGCTCTGTTTCCACCGGGTACACCAAGTGAAT 1440
DB 1381 ACCATTCTCAGTATCTGACCAAGGCTCTGTTTCCACCGGGTACACCAAGTGAAT 1440
QY 1441 GCATTGTCAAAGCAGCTGCAATGCTGGAATAATATATGAGGGCTTGTGATTTGGC 1500
DB 1441 GCATTGTCAAAGCAGCTGCAATGCTGGAATAATATATGAGGGCTTGTGATTTGGC 1500
QY 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533
DB 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533

RESULT 3
ADQ14504
ID ADQ14504 standard; cDNA; 1533 BP.

AC ADQ14504;

XX 23-SEP-2004 (first entry)

DE wild type soybean myo-inositol 1-phosphate synthase cDNA #2.

XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate.

XX Glycine max.

XX Key Location/Qualifiers

XX CDS 1..1533

FT /product= "a" wild type soybean myo-inositol 1-phosphate
FT /synthase #2"

XX US2004128713-A1.

XX PD 01-JUL-2004.

XX PF 21-NOV-2003; 2003US-00718952.

XX PR 08-APR-1997; 97US-00835751.

XX PR 07-APR-1998; 98WO-US0068822.

XX PR 26-APR-1999; 99US-00299315.

XX PR 11-MAR-2002; 2002US-00025003.

XX PA (HITZ/) HITZ W D.

XX PA (SEBA/) SEBASTIAN S A.

XX PA (GRAC/) GRACE D J.

XX PA (STRE/) STREIT L G.

XX PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI; 2004-53135/51.
DR P-PSDB; ADQ14505.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX Claim 4; SEQ ID NO 15; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 96.7%; Score 1482; DB 12; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTCATCGAGATTTTAAAGTAGAGAGTCTTAATGTAAGTACCGGAGCTAGATT 60
DB 1 ATGTCATCGAGATTTTAAAGTAGAGAGTCTTAATGTAAGTACCGGAGCTAGATT 60
QY 61 CAGTCGGGTACAACTAGAAACCGAAGCTGTTGTCAGAGAACGAAATGGACCTAT 120
DB 61 CAGTCGGGTACAACTAGAAACCGAAGCTGTTGTCAGAGAACGAAATGGACCTAT 120
QY 121 CAGTGAATTTGCAAAACCAATCCGTCACCAATTTAAACCAACCAACCATTTTCA 180
DB 121 CAGTGAATTTGCAAAACCAATCCGTCACCAATTTAAACCAACCAACCATTTTCA 180
QY 181 AAATTGGGGGTGATCTTGTGGGTGGGTAACAACGCTCTACCTCACCGGTGT 240
DB 181 AAATTGGGGGTGATCTTGTGGGTGGGTAACAACGCTCTACCTCACCGGTGT 240
QY 241 GTTATTGCTAACAGAGAGACATTTATGAGGCTAACAAAGACCAAGTTCACCAAT 300
DB 241 GTTATTGCTAACAGAGAGACATTTATGAGGCTAACAAAGACCAAGTTCACCAAT 300
QY 301 TACTTTGGCTCCCTCACCAAGCCTCAGCTATTGAGTTGATCTTTCAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCACCAAGCCTCAGCTATTGAGTTGATCTTTCAGGAGAGAA 360
QY 361 ATCTATGCCCATTCAGAGATGCTGCTTCAATGTTAATCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCATTCAGAGATGCTGCTTCAATGTTAATCTGACGACATTTGTTGGG 420
QY 421 GGATGGGATATCAGAAATGAACCTGCTGATGCAAGGCGCAAGGAGGATGTTGAC 480
DB 421 GGATGGGATATCAGAAATGAACCTGCTGATGCAAGGCGCAAGGAGGATGTTGAC 480
QY 481 ATGATTTGCAAGACAGATTGAGGCTTTACATGATCATGATTTTCACTCCCGGAATC 540
DB 481 ATGATTTGCAAGACAGATTGAGGCTTTACATGATCATGATTTTCACTCCCGGAATC 540
QY 541 TAGACCCGGATTTTCAATTGCTGCCAACCAAGAGAGCGTCCCAACAGCTGTTAAGGCG 600

```
Db 541 TAGACCCGATTTCTATGCTGCAACAGAGAGCGTCCCAACACGGATTAAGGCG 600
Qy 601 ACAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGCGCTTTAAGAGCCACC 660
Db 601 ACAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGCGCTTTAAGAGCCACC 660
Qy 661 AAGTGGACAGGTTGTTGCTGCTGAGACTGCAACAGAGAGGATATGCAATTTGGTT 720
Db 661 AAGTGGACAGGTTGTTGCTGCTGAGACTGCAACAGAGAGGATATGCAATTTGGTT 720
Qy 721 GTAGGCTTAATGACACCAATGAGAAATCTTTGGCTGTGTGAGCAAGAAATGAGCGTGA 780
Db 721 GTAGGCTTAATGACACCAATGAGAAATCTTTGGCTGTGTGAGCAAGAAATGAGCGTGA 780
Qy 781 ATTTCTCTTCCACCTGTTATGTCATGCTGCTGTGTATGAGAAATGTTCTTTCAATTA 840
Db 781 ATTTCTCTTCCACCTGTTATGTCATGCTGCTGTGTATGAGAAATGTTCTTTCAATTA 840
Qy 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATTTGCCATCGGAGAACACT 900
Db 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATTTGCCATCGGAGAACACT 900
Qy 901 TTGATTTGAGAGATGACTTCAAGATGCTCAGACCAACCAATCTGTGTTGGTTGAT 960
Db 901 TTGATTTGAGAGATGACTTCAAGATGCTCAGACCAACCAATCTGTGTTGGTTGAT 960
Qy 961 TTTCTTGTGGGGCTGTGTATCAAGCCCAACTATATAGTTAGTTACCAACCACTGGGAA 1020
Db 961 TTTCTTGTGGGGCTGTGTATCAAGCCCAACTATATAGTTAGTTACCAACCACTGGGAA 1020
Qy 1021 AATGATGATGATGATTTCTTGGCTCCCAACAACTTCCGCTCCAGAGAAATCTTCAAGAG 1080
Db 1021 AATGATGATGATGATTTCTTGGCTCCCAACAACTTCCGCTCCAGAGAAATCTTCAAGAG 1080
Qy 1081 AACGTTGTGAGCAATGAGTCAACAGCAATGCAATGCTCTATATAGCTGTGTAACATCCC 1140
Db 1081 AACGTTGTGAGCAATGAGTCAACAGCAATGCAATGCTCTATATAGCTGTGTAACATCCC 1140
Qy 1141 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAAGGGATATGACAAAGAGCCATGAT 1200
Db 1141 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAAGGGATATGACAAAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAGAGATTTTCATGCGTGGAAAAAACCATTGTTTTCACACACATGT 1260
Db 1201 GAGTACACTTCAGAGATTTTCATGCGTGGAAAAAACCATTGTTTTCACACACATGT 1260
Qy 1261 GAGGATTCCTTTAGCTGCTCTATATCTTGGACTTGGCTTCTGCTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTAGCTGCTCTATATCTTGGACTTGGCTTCTGCTGAGCTGAGC 1320
Qy 1321 ACTAGAAATCAGTTTAAAGCTGAAAAATGAGGAAAAATTCACACTTTCACCCAGTTGCT 1380
Db 1321 ACTAGAAATCAGTTTAAAGCTGAAAAATGAGGAAAAATTCACACTTTCACCCAGTTGCT 1380
Qy 1381 ACCATTCTAGCTATTTGACCAAGGCTCCTTGTGTTCCACCGGGTACACAGTGTGAAT 1440
Db 1381 ACCATTCTAGCTATTTGACCAAGGCTCCTTGTGTTCCACCGGGTACACAGTGTGAAT 1440
Qy 1441 GCATTGTCAAGGAGGTGCAATGCTGAAAAACATATAGAGGCTGTGTGATTTGGCC 1500
Db 1441 GCATTGTCAAGGAGGTGCAATGCTGAAAAACATATAGAGGCTGTGTGATTTGGCC 1500
Qy 1501 CCAGAGAAATACATGATTTCTCGAGTACAACTGA 1533
Db 1501 CCAGAGAAATACATGATTTCTCGAGTACAACTGA 1533
```

```
RESULT 4
ADQ14502
ID ADQ14502 standard; cDNA, 1533 BP.
XX
AC ADQ14502;
```

```
XX 23-SEP-2004 (first entry)
DT
XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
DE
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
XX myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.
XX
XX glycine max.
OS
XX Synthetic.
FH
XX Key Location/Qualifiers
FT 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #4"
PN
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US0006822.
XX 26-APR-1999; 99US-0029315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI: 2004-53335/51.
XX P-PSDB; ADQ14503.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Example 8; SEQ ID NO 13; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents cDNA encoding a mutant soybean
XX myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 96.7%; Score 1482; DB 12; Length 1533;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 ATGTTCAATGAGAAATTTAAGGTAGAGAGTCTTAATGTAAGTACGAGAGCTGAGATT 60
```

Db 1 ATGTTTCATCGAAGATTTTAAGGTAGAGAGTCTCTAATGTGAATGACACCGAGCTGAGATT 60
Qy 61 CAGTCCGTGTACAACTACGAAACCAACCGAATTGTTTCAAGAGACAGAAATGGACCTAT 120
Db 61 CAGTCCGTGTACAACTACGAAACCAACCGAATTGTTTCAAGAGACAGAAATGGACCTAT 120
Qy 121 CAGTGGATTGTCAAAACCCAAATCCGTCAACTACCAATTTAAAAACAACACCCATGTTCCA 180
Db 121 CAGTGGATTGTCAAAACCCAAATCCGTCAACTACCAATTTAAAAACAACACCCATGTTCCA 180
Qy 181 AAATTGGGGGTGATGCTGTGGGTGGGTTGGGTTGAAACAACGGCTTACCTGACCGGTGT 240
Db 181 AAATTGGGGGTGATGCTGTGGGTGGGTTGAAACAACGGCTTACCTGACCGGTGT 240
Qy 241 GTTATTGTCTACAGAGAGACATTTATGAGGCTACAAAGAGACAAAGATTCAACAGCCAT 300
Db 241 GTTATTGTCTACAGAGAGAGACATTTATGAGGCTACAAAGAGACAAAGATTCAACAGCCAT 300
Qy 301 TACTTTGGCTCCCTGACCCCAAGCCTCAAGCTATTCGAGTTGATCCTTCCAGGGAGAGAA 360
Db 301 TACTTTGGCTCCCTGACCCCAAGCCTCAAGCTATTCGAGTTGATCCTTCCAGGGAGAGAA 360
Qy 361 ATCTATGCCCCATTCAAGAGTCTGCTTCCATGAGTTAACTCTGACGACATTTGTTGGG 420
Db 361 ATCTATGCCCCATTCAAGAGTCTGCTTCCATGAGTTAACTCTGACGACATTTGTTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGTCATGGCCAGGGCAAAAGGTTTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGTCATGGCCAGGGCAAAAGGTTTGAC 480
Qy 481 ATGATTTTGGCAGAGACAGTTGAGGCTTTACATGAATTCATGAGTTTCCATCCCGGAATC 540
Db 481 ATGATTTTGGCAGAGACAGTTGAGGCTTTACATGAATTCATGAGTTTCCATCCCGGAATC 540
Qy 541 TACGACCCCGGATTTTATGCTGCTCCAAACGAGAGGCGTGCCAAACCTGATTAAGAGGC 600
Db 541 TACGACCCCGGATTTTATGCTGCTCCAAACGAGAGGCGTGCCAAACCTGATTAAGAGGC 600
Qy 601 ACAAGACAGAGACAGTTGAGCAAAATCATCAAGACATCAAGGCTTTTAAGAACCCACC 660
Db 601 ACAAGACAGAGACAGTTGAGCAAAATCATCAAGACATCAAGGCTTTTAAGAACCCACC 660
Qy 661 AAAGTGAACAGGTGGTGTCTCTGTGACTGCCAACACAGAGAGTATGCAATTTGGTT 720
Db 661 AAAGTGAACAGGTGGTGTCTCTGTGACTGCCAACACAGAGAGTATGCAATTTGGTT 720
Qy 721 GTAGGCTTTAATGACACCAATGGAATCTTGGCTGTGTGACAGAAATGAGGCTGAG 780
Db 721 GTAGGCTTTAATGACACCAATGGAATCTTGGCTGTGTGACAGAAATGAGGCTGAG 780
Qy 781 ATTTCTCCTTCCACCTGTGATGCCATTTGCTGTGTGATGGAATAATGTTCTTTCAATAT 840
Db 781 ATTTCTCCTTCCACCTGTGATGCCATTTGCTGTGTGATGGAATAATGTTCTTTCAATAT 840
Qy 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATTTGATCTTTGCAATGCGAGAACT 900
Db 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATTTGATCTTTGCAATGCGAGAACT 900
Qy 901 TTGATTGTGGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTGAT 960
Db 901 TTGATTGTGGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTGAT 960
Qy 961 TTTCTTGTGGGGGTGTGTATCAAGCAATCATATAGTTAGTTAACACCATGAGGAAC 1020
Db 961 TTTCTTGTGGGGGTGTGTATCAAGCAATCATATAGTTAGTTAACACCATGAGGAAC 1020
Qy 1021 AATGATGGTATGATCTCTGGCTCCCAAAACCTTCCGCTCAAGAGAAATCTCAAGAC 1080
Db 1021 AATGATGGTATGATCTCTGGCTCCCAAAACCTTCCGCTCAAGAGAAATCTCAAGAC 1080
Qy 1081 AACGTTTGTGAGATGATGTCACAGCAATGCCATCTCTATAGAGCTGTGTGAACATCC 1140
Db 1081 AACGTTTGTGAGATGATGTCACAGCAATGCCATCTCTATAGAGCTGTGTGAACATCC 1140

Qy 1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGATAGCAAGAGCCATGGAT 1200
Db 1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGATAGCAAGAGCCATGGAT 1200
Qy 1201 GAGTACATCTTCAAGATTTTATGAGGTGGAAGAACACCAATTTGTTTGCACAACATATG 1260
Db 1201 GAGTACATCTTCAAGATTTTATGAGGTGGAAGAACACCAATTTGTTTGCACAACATATG 1260
Qy 1261 GAGGATTCCTTTTATGAGTGTCTCTATTAATCTTGGACCTTGCTCTTGTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTTATGAGTGTCTCTATTAATCTTGGACCTTGCTCTTGTGAGCTGAGC 1320
Qy 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCACATATTCACCCAGTTGCT 1380
Db 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCACATATTCACCCAGTTGCT 1380
Qy 1381 ACCATTTCTGAGCTATCTGACCAAGGCTCTCTGTTCCACCGGGTACACAGTGTGAT 1440
Db 1381 ACCATTTCTGAGCTATCTGACCAAGGCTCTCTGTTCCACCGGGTACACAGTGTGAT 1440
Qy 1441 GCATTGTCAAGACGCGTGAATGCTGGAATAACATATGAGGCTTGTGATTGAGCC 1500
Db 1441 GCATTGTCAAGACGCGTGAATGCTGGAATAACATATGAGGCTTGTGATTGAGCC 1500
Qy 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533
Db 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 5
ADS82005
ID ADS82005 standard; cDNA; 1533 BP.
XX
XX
AC ADS82005;
XX
DT 18-NOV-2004 (first entry)
XX
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
XX
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ssi; gene; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag=a
FT /product="myo-inositol 1-phosphate synthase"
FT
XX
XX
PN US2003074685-A1.
XX
XX
PD 17-APR-2003.
XX
XX
PF 11-MAR-2002; 2002US-00025003.
XX
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX
PI Hitz WD, Sebastian SA;
XX
XX
DR MPI. 2004-63957/62.
DR P-PSDB; ADS82006.
XX
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX
PS Claim 2, SEQ ID NO 15; 34bp; English.
XX

The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype with (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 96.7%; Score 1482; DB 13; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGTTATGAGAAATTTAAAGTAGAGAGCTTAATGTGAATCAACCGAGATGAGATT 60
1 ATGTTATGAGAAATTTAAAGTAGAGAGCTTAATGTGAATCAACCGAGATGAGATT 60
61 CAGTCCGCTGACAACTACGAAACCAACGAACTGTTACAGAGAAAGAAATGACCTAT 120
61 CAGTCCGCTGACAACTACGAAACCAACGAACTGTTACAGAGAAAGAAATGACCTAT 120
121 CAGTGGATTGTCAAAACCAATCGTCACTACCAATTTAAACCAACCAATGTTTCA 180
121 CAGTGGATTGTCAAAACCAATCGTCACTACCAATTTAAACCAACCAATGTTTCA 180
181 AAATTTGGGGTGTATGCTTGTGGTGGTGGTGAACAAGGCTCTACCGTGGT 240
181 AAATTTGGGGTGTATGCTTGTGGTGGTGGTGAACAAGGCTCTACCGTGGT 240
241 GTTATTTGCTAACAGAGAGACATTTATGAGGCTACAAAGAGCAAGATTCACAGCAAT 300
241 GTTATTTGCTAACAGAGAGACATTTATGAGGCTACAAAGAGCAAGATTCACAGCAAT 300
301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTTCCAGGAGAGAA 360
301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTTCCAGGAGAGAA 360
361 ATCTATGAGCCCAATTAAAGAGTCTGCTTCAATGTTAACTCGACAGATGTTGGG 420
361 ATCTATGAGCCCAATTAAAGAGTCTGCTTCAATGTTAACTCGACAGATGTTGGG 420
421 GGATGGAGATATCAGCAAGAGAACTGCTGATGATGCGCAGGCAAGAGTGTTCAC 480
421 GGATGGAGATATCAGCAAGAGAACTGCTGATGATGCGCAGGCAAGAGTGTTCAC 480
481 ATGATTTGCAAGAGAGAGTTGAGGCTTACATGAAATTCATGTTCCACTCCCGAATC 540
481 ATGATTTGCAAGAGAGAGTTGAGGCTTACATGAAATTCATGTTCCACTCCCGAATC 540

541 TACGACCCGGAATTTATGCTGCAACCAAGAGAGAGCTGCCAACAACGATTAAGGC 600
541 TACGACCCGGAATTTATGCTGCAACCAAGAGAGAGCTGCCAACAACGATTAAGGC 600
601 ACAAGCAAG 660
601 ACAAGCAAG 660
661 AAAGTGAAG 720
661 AAAGTGAAG 720
721 GTAGGCTTAATGAG 780
721 GTAGGCTTAATGAG 780
781 ATTTCTCTTCCACCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
781 ATTTCTCTTCCACCTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
841 GGAAGCCCTCAG 900
841 GGAAGCCCTCAG 900
901 TTGATTGAG 960
901 TTGATTGAG 960
961 TTTCTGTGGGGCTGCTGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
961 TTTCTGTGGGGCTGCTGTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
1021 AATGATGATATGATATCTCTGCTCCACCAACCTTCCTCCAGAGAAATCTCCAGAG 1080
1021 AATGATGATATGATATCTCTGCTCCACCAACCTTCCTCCAGAGAAATCTCCAGAG 1080
1081 AAGCTTGTGAGATATGATATGATATGATATGATATGATATGATATGATATGATAT 1140
1081 AAGCTTGTGAGATATGATATGATATGATATGATATGATATGATATGATATGATAT 1140
1141 GACATGATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1200
1141 GACATGATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 1200
1201 GAGTACACTTCAAGAGATATGATATGATATGATATGATATGATATGATATGATAT 1260
1201 GAGTACACTTCAAGAGATATGATATGATATGATATGATATGATATGATATGATAT 1260
1261 GAGATTCCTTTTATGCTGCTCTATATCTTGAAGAGAGAGAGAGAGAGAGAGAGAG 1320
1261 GAGATTCCTTTTATGCTGCTCTATATCTTGAAGAGAGAGAGAGAGAGAGAGAGAG 1320
1321 ACTAGATCACTTTAAAG 1380
1321 ACTAGATCACTTTAAAG 1380
1381 ACCATTTCTAGATATGATATGATATGATATGATATGATATGATATGATATGATAT 1440
1381 ACCATTTCTAGATATGATATGATATGATATGATATGATATGATATGATATGATAT 1440
1441 GGAATGCAAG 1500
1441 GGAATGCAAG 1500
1501 CCAAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 1533
1501 CCAAGAGATATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 1533

RESULT 6
AD582003
ID AD582003 standard; cDNA; 1533 BP.
XX

AC ADS82003;
XX 18-NOV-2004 (first entry)
DT
XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
DE
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29018JP03.
XX
FH Key Location/Qualifiers
FT 1..1533
FT CDS /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-63957/62.
XX P-PSDB; ADS82004.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phylic acid and inorganic phosphate content of soybean seeds.
XX
XX Example 8; SEQ ID NO 13; 34bp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phylic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homologous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate), where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phylic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
XX inositol 1-phosphate synthase.
XX
SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 96.7%; Score 1482; DB 13; Length 1533;

		Best Local Similarity 99.9%; Pred. No. 0;				
		Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
Qy	1	ATGTCATCGAGAAATTTTAAAGTAGAGAGCTCTAATGTGAAGTACACCGAGCTGAGATT				60
Db	1	ATGTCATCGAGAAATTTTAAAGTAGAGAGCTCTAATGTGAAGTACACCGAGCTGAGATT				60
Qy	61	CAGTCCTGTATCAACTAGCAAAACCCGAACTTGTTCAGGAACAGGAATGGACCTAT				120
Db	61	CAGTCCTGTATCAACTAGCAAAACCCGAACTTGTTCAGGAACAGGAATGGACCTAT				120
Qy	121	CAGTCGATTTGCAAAACCCGAAATCCGTCAACTAGCAATTTTAAACCAACCACTGTTCA				180
Db	121	CAGTCGATTTGCAAAACCCGAAATCCGTCAACTAGCAATTTTAAACCAACCACTGTTCA				180
Qy	181	AAATGGGGGTGATGCTTGGGGTGGGGTGGAAACCAAGGCTCTACCTGACCGGTGT				240
Db	181	AAATGGGGGTGATGCTTGGGGTGGGGTGGAAACCAAGGCTCTACCTGACCGGTGT				240
Qy	241	GTTATTTGCTAACAGAGAGACATTTGATGGCTACAAAGACAGATTCACCAAGCCAT				300
Db	241	GTTATTTGCTAACAGAGAGAGACATTTGATGGCTACAAAGACAGATTCACCAAGCCAT				300
Qy	301	TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCGAGTTGGATCTTCCAGGAGAGAA				360
Db	301	TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCGAGTTGGATCTTCCAGGAGAGAA				360
Qy	361	ATCTATGCCCCATTCAAGAGTCTGCTTCAATGTTTAACTCGACGACATTTGTTGG				420
Db	361	ATCTATGCCCCATTCAAGAGTCTGCTTCAATGTTTAACTCGACGACATTTGTTGG				420
Qy	421	GGATGGATATCAGCAACATGAACCTGGCTGATGTCATGCGCCAGGCAAGGTGTTGAC				480
Db	421	GGATGGATATCAGCAACATGAACCTGGCTGATGTCATGCGCCAGGCAAGGTGTTGAC				480
Qy	481	ATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC				540
Db	481	ATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC				540
Qy	541	TACGACCCGGAATTTATGCTGACCAACCAAGAGAGCGTGCACAAAGTATTAAAGGC				600
Db	541	TACGACCCGGAATTTATGCTGACCAACCAAGAGAGCGTGCACAAAGTATTAAAGGC				600
Qy	601	ACAAAGCAAGAGCAAGTTGAGCAATTCATCAAGACATCAAGGCTTTAAAGAACCAAC				660
Db	601	ACAAAGCAAGAGCAAGTTGAGCAATTCATCAAGACATCAAGGCTTTAAAGAACCAAC				660
Qy	661	AAAGTGACAAAGTGTGTTCTCTGTGACCTGCCAACAGAGAGGTATAGCAATTTGGTT				720
Db	661	AAAGTGACAAAGTGTGTTCTCTGTGACCTGCCAACAGAGAGGTATAGCAATTTGGTT				720
Qy	721	GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG				780
Db	721	GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG				780
Qy	781	ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGATGAAATGTTCTTTCAATAT				840
Db	781	ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGATGAAATGTTCTTTCAATAT				840
Qy	841	GGAACCCCTCAGAAACATTTTGTACAGGCTGATTTGATCTTTGCCATGGCAGAACT				900
Db	841	GGAACCCCTCAGAAACATTTTGTACAGGCTGATTTGATCTTTGCCATGGCAGAACT				900
Qy	901	TTGATTTGGGATGACCTTCAAGAGTGTGACCAAAATGAATCTGTGTGTTGAT				960
Db	901	TTGATTTGGGATGACCTTCAAGAGTGTGACCAAAATGAATCTGTGTGTTGAT				960
Qy	961	TTTCTTGTGGGGCTGTATCAAGCCAACTATAGTTAGTTACACCATCTGGGAAAC				1020
Db	961	TTTCTTGTGGGGCTGTATCAAGCCAACTATAGTTAGTTACACCATCTGGGAAAC				1020
Qy	1021	AATGATGATGATATCTCGGCTCCAAACCTTCGCTCCAGGAAATCTCCAAGAC				1080
Db	1021	AATGATGATGATATCTCGGCTCCAAACCTTCGCTCCAGGAAATCTCCAAGAC				1080

Db 1021 AATGATGATGATGATCTCTGGCTGCCAACAACCTTCGCTCCAGAAAGAAATCTCCAGAGC 1080
 Qy 1081 AAGTTGTTGACATATGCTCAACAGAAAGCCATCTCTATGAGCTGTGTGAACATCCC 1140
 Db 1081 AAGTTGTTGACATATGCTCAACAGAAAGCCATCTCTATGAGCTGTGTGAACATCCC 1140
 Qy 1141 GACCATGTTGTTGTTAATGATGCTTACGTTAGGGGATGCAAGAGCCATGGAT 1200
 Db 1141 GACCATGTTGTTGTTAATGATGCTTACGTTAGGGGATGCAAGAGCCATGGAT 1200
 Qy 1201 GACTACACTTCAGAGATATTCATGGGTGGAAGAACACCATTTGTTGCAACACATGT 1260
 Db 1201 GACTACACTTCAGAGATATTCATGGGTGGAAGAACACCATTTGTTGCAACACATGT 1260
 Qy 1261 GAGGATTCCTTTTACGTCCTCTATTTATCTTGAAGCTGGTCTTTGTTGAGCTGAGC 1320
 Db 1261 GAGGATTCCTTTTACGTCCTCTATTTATCTTGAAGCTGGTCTTTGTTGAGCTGAGC 1320
 Qy 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACATCTCCAGCTGCT 1380
 Db 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACATCTCCAGCTGCT 1380
 Qy 1381 ACCATCTCAGCTATCTGACCAAGGCTCTCTGTTCCACCGGTTACACCAAGTGTGAAT 1440
 Db 1381 ACCATCTCAGCTATCTGACCAAGGCTCTCTGTTCCACCGGTTACACCAAGTGTGAAT 1440
 Qy 1441 GCATTTGCAAGAGCGGTGCAATGCTGGAAGAAACATATGAGGGCTGTGTGATTTGGCC 1500
 Db 1441 GCATTTGCAAGAGCGGTGCAATGCTGGAAGAAACATATGAGGGCTGTGTGATTTGGCC 1500
 Qy 1501 CCAGAGATATACATGATTTCTCGAGTACAGTGA 1533
 Db 1501 CCAGAGATATACATGATTTCTCGAGTACAGTGA 1533

RESULT 7
 AAV62443
 ID AAV62443 standard; cDNA; 1533 BP.

AC AAV62443;
 AT 17-OCT-2003 (revised)
 DT 02-FEB-1999 (first entry)
 XX
 DE Soybean mutant myo-inositol 1-phosphate synthase cDNA.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
 KM phytic acid; dg.
 XX
 OS Glycine max; line LR33.
 XX
 PN W09845448-A1.
 XX
 PD 15-OCT-1998.
 XX
 PR 07-APR-1998; 98WO-US006822.
 XX
 PR 08-APR-1997; 97US-00835751.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Hitz WD, Sebastian SA;
 XX
 DR WPI, 1998-568353/48.
 DR P-PSDB; AAW79741.
 XX
 PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
 PT for generating plants with altered levels of e.g. raffinose, stachyose,
 PT phytic acid, etc.
 XX
 PS Example 5; Page 48-49; 63pp; English.
 CC This is the nucleotide sequence of cDNA encoding a mutant soybean myo-

CC inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
 CC clone was isolated from a cDNA library of soybean line LR33 by PCR
 CC amplification (see AAV62441-42). Line LR33 was obtained by chemical
 CC mutagenesis of wild-type soybean genome and as a reduced raffinose
 CC saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
 CC acid, raffinose and stachyose. Sequencing revealed a single base change
 CC mutation (G to T at base 1241) in the LR33 sequence when compared to the
 CC wild-type sequence (see AAV62440). The mutation results in a seed
 CC phenotype of very low raffinose saccharide sugars, very high sucrose and
 CC low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
 CC raffinose saccharide, sucrose, phytic acid and inorganic phosphate
 CC content of soybean seeds, leading to useful soybean products, e.g. a seed
 CC phytic acid content of less than 17 ug/g, a seed content of raffinose and
 CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
 CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 9.5%; Score 146; DB 2; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 1e-60; Mismatches 0; Gaps 0;

Db 817 ATGAAATGTTCTTTCAATATGGAAGCCCTCAGAACATTTGTACAGGGCTGAT 876
 Qy 817 ATGAAATGTTCTTTCAATATGGAAGCCCTCAGAACATTTGTACAGGGCTGAT 876
 Db 877 GATCTTGCATGCGAGAACACTTTGATTTGATGATGACTTCAAGAGTGTCAAGC 936
 Qy 877 GATCTTGCATGCGAGAACACTTTGATTTGATGATGACTTCAAGAGTGTCAAGC 936
 Db 937 AAAATGAATCTGTGTGTTGATTT 962
 Qy 937 AAAATGAATCTGTGTGTTGATTT 962
 Db 937 AAAATGAATCTGTGTGTTGATTT 962

RESULT 8
 ADQ14498
 ID ADQ14498 standard; cDNA; 1533 BP.

AC ADQ14498;
 AT 23-SEP-2004 (first entry)
 DT
 XX
 DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 XX inorganic phosphate; mutant.
 XX
 OS Glycine max.
 XX
 PN Synthetic.
 XX
 FH Key Location/Qualifiers
 FT 1. 1533
 FT CDS /tag= a
 FT /product= "Mutant soybean myo-inositol 1-phosphate
 FT synthase #2"
 XX
 PN US2004128713-A1.
 XX
 PD 01-JUL-2004.
 XX
 PF 21-NOV-2003; 2003US-00718952.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-00299315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STREIT/) STREIT L G.


```
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
PI WPI: 2004-53135/51.
XX P-PSDB; ADQ14499.
DR
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX Example 8; SEQ ID NO 9; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
Query Match
Best Local Similarity 9.5%; Score 146; DB 12; Length 1533;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 817 ATGGAATAATGTTCTTTCATTATATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
DB 817 ATGGAATAATGTTCTTTCATTATATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
QY 877 GATCTTGGCATTGCGAGAGAACCTTTGATTGGTGAGATGACTTCAAGAGTGCTCAGACC 936
DB 877 GATCTTGGCATTGCGAGAGAACCTTTGATTGGTGAGATGACTTCAAGAGTGCTCAGACC 936
QY 937 AAAATGAATCTGTGTGGTTGATT 962
DB 937 AAAATGAATCTGTGTGGTTGATT 962
RESULT 9
ADQ14494
ID ADQ14494 standard; cDNA; 1533 BP.
XX
XX ADQ14494;
XX
XX 23-SEP-2004 (first entry)
XX
XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.
XX
XX Glycine max.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /*tag= a
XX FT /product= "Mutant soybean myo-inositol 1-phosphate
XX FT synthase #1"
```

```
FT mutation replace(1241,G)
XX /*tag= b
XX
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006882.
XX 26-APR-1999; 99US-00289315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
PI WPI: 2004-53135/51.
XX P-PSDB; ADQ14495.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Claim 10; SEQ ID NO 5; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
Query Match
Best Local Similarity 9.5%; Score 146; DB 12; Length 1533;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 817 ATGGAATAATGTTCTTTCATTATATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
DB 817 ATGGAATAATGTTCTTTCATTATATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
QY 877 GATCTTGGCATTGCGAGAGAACCTTTGATTGGTGAGATGACTTCAAGAGTGCTCAGACC 936
DB 877 GATCTTGGCATTGCGAGAGAACCTTTGATTGGTGAGATGACTTCAAGAGTGCTCAGACC 936
QY 937 AAAATGAATCTGTGTGGTTGATT 962
DB 937 AAAATGAATCTGTGTGGTTGATT 962
RESULT 10
ADQ81999
ID ADQ81999 standard; cDNA; 1533 BP.
XX
```


AC ADS81999;
 XX 18-NOV-2004 (first entry)
 XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
 XX
 DE Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 KM
 XX Glycine max; line 29004DP01.
 OS
 XX Key Location/Qualifiers
 FH 1..1533
 FT CDS /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 XX
 XX US2003074685-A1.
 XX
 XX 17-APR-2003.
 XX
 XX 11-MAR-2002; 2002US-00025003.
 XX
 XX 08-APR-1997; 97US-00835751.
 XX 07-APR-1998; 98WO-US006822.
 XX
 XX (HITZ/) HITZ W D.
 XX (SEBA/) SEBASTIAN S A.
 XX
 XX Hitz WD, Sebastian SA;
 XX
 XX WPI: 2004-639957/62.
 XX P-PSDB; ADS82000.
 DR
 XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 XX
 XX Example 8; SEQ ID NO 9; 34pp; English.
 PS
 XX The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with a
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence encodes a wild-type myo-
 CC inositol 1-phosphate synthase.
 CC
 CC Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
 CC
 CC Query Match 9.5%; Score 146; DB 13; Length 1533;

Best Local Similarity 100.0%; Pred. No. 1e-60;
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 817 ATGAAATGCTCTTCAATATGAGACCCCTCAGACACTTTGTACAGGCTGATT 876
 DB 817 ATGAAATGCTCTTCAATATGAGACCCCTCAGACACTTTGTACAGGCTGATT 876
 QY 877 GATCTTGCATGCGCAGGAACCTTGTATGAGATGATCTCAAGTGTGACACC 936
 DB 877 GATCTTGCATGCGCAGGAACCTTGTATGAGATGATCTCAAGTGTGACACC 936
 QY 937 AAAATGAATCTGTGTTGTTGATT 962
 DB 937 AAAATGAATCTGTGTTGTTGATT 962
 RESULT 11
 ADS81993
 ID ADS81993 standard; cDNA; 1533 BP.
 XX
 XX ADS81993;
 XX
 XX 18-NOV-2004 (first entry)
 XX
 XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
 XX
 XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 XX
 XX Glycine max; cultivar Wye.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1..1533
 FT CDS /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 XX
 XX US2003074685-A1.
 XX
 XX 17-APR-2003.
 XX
 XX 11-MAR-2002; 2002US-00025003.
 XX
 XX 08-APR-1997; 97US-00835751.
 XX 07-APR-1998; 98WO-US006822.
 XX
 XX (HITZ/) HITZ W D.
 XX (SEBA/) SEBASTIAN S A.
 XX
 XX Hitz WD, Sebastian SA;
 XX
 XX WPI: 2004-639957/62.
 XX P-PSDB; ADS81994.
 DR
 XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 XX
 XX Claim 2; SEQ ID NO 1; 34pp; English.
 PS
 XX The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype

(comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Query Match 9.5%; Score 146; DB 13; Length 1533;
Best Local Similarity 100.0%; Pred. No. 1e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

817 ATGGAAGAAATGTTCTTTCATTATGGAACCCCTCAGAAACATTGTATCCAGGCTGATT 876
|||||
817 ATGGAAGAAATGTTCTTTCATTATGGAACCCCTCAGAAACATTGTATCCAGGCTGATT 876

877 GATCTTGCCATCGCAGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGACC 936
|||||
877 GATCTTGCCATCGCAGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGACC 936

937 AAAATGAAATCTGTGTGTTGATT 962
|||||
937 AAAATGAAATCTGTGTGTTGATT 962

RESULT 12
ADS81997
ID ADS81997 standard; cDNA; 1533 BP.
AC ADS81997;
XX
XX 18-NOV-2004 (first entry)
DT
DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
OS Glycine max; line LR33.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH CDS 1..1533
FT /*tag= a
FT /product= "myo-inositol 1-phosphate synthase"
FT replace(1188,G)
FT /*tag= b
XX mutation
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX

WPI, 2004-639957/62.
DR P-PSDB; ADS81998.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 5; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a mutant myo-inositol 1-phosphate synthase.

Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 9.5%; Score 146; DB 13; Length 1533;
Best Local Similarity 100.0%; Pred. No. 1e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

817 ATGGAAGAAATGTTCTTTCATTATGGAACCCCTCAGAAACACTTTGTACAGGCTGATT 876
|||||
817 ATGGAAGAAATGTTCTTTCATTATGGAACCCCTCAGAAACACTTTGTACAGGCTGATT 876

877 GATCTTGCCATCGCAGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGACC 936
|||||
877 GATCTTGCCATCGCAGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGACC 936

937 AAAATGAAATCTGTGTGTTGATT 962
|||||
937 AAAATGAAATCTGTGTGTTGATT 962

RESULT 13
ADQ14490
ID ADQ14490 standard; cDNA; 1760 BP.
AC ADQ14490;
XX
XX 23-SEP-2004 (first entry)
DT
DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #1.
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate.
XX
XX

OS Glycine max.
 XX
 FH Key Location/Qualifiers
 FT CDS 54..1586
 FT /tag= a
 FT /product= "Wild type soybean myo-inositol 1-phosphate
 FT synthase #1"
 XX
 XX US2004128713-A1.
 XX
 XX
 XX
 XX PD 01-JUL-2004.
 XX
 XX PF 21-NOV-2003; 2003US-00718952.
 XX
 XX PR 08-APR-1997; 97US-00835751.
 XX PR 07-APR-1998; 98WO-US006822.
 XX PR 26-APR-1999; 99US-00293315.
 XX PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 XX
 XX WPI; 2004-533135/51.
 XX P-PSDB; ADQ14491.
 XX
 PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX
 XX Claim 4; SEQ ID NO 1; 48pp; English.
 XX
 XX The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant
 CC seed or the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents cDNA encoding a wild type
 CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
 XX
 XX Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;
 SQ
 Query Match 9.5%; Score 146; DB 12; Length 1760;
 Best Local Similarity 100.0%; Pred. NO. 1e-60;
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
 AAV62440.
 ID AAV62440 standard; cDNA; 1782 BP.
 XX
 XX AAV62440;
 AC
 XX 17-OCT-2003 (revised)
 DT 02-FEB-1999 (first entry)
 XX
 XX Soybean wild-type myo-inositol 1-phosphate synthase cDNA.
 DE
 XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
 XX phytic acid; ds.
 XX
 XX Glycine max; line LR13.
 OS
 XX
 XX
 XX
 XX
 FT Key Location/Qualifiers
 FT CDS 54..1586
 FT /tag= a
 XX
 XX W09845448-A1.
 XX
 XX PD 15-OCT-1998.
 XX
 XX PF 07-APR-1998; 98WO-US006822.
 XX
 XX PR 08-APR-1997; 97US-00835751.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 XX Hitz WD, Sebastian SA;
 PI
 XX
 XX WPI; 1998-568353/48.
 XX P-PSDB; AAW79740.
 XX
 PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
 PT for generating plants with altered levels of e.g. raffinose, stachyose,
 PT phytic acid, etc.
 XX
 XX Example 5; Page 44-45; 63pp; English.
 XX
 XX This is the nucleotide sequence of cDNA encoding the wild-type soybean
 CC myo-inositol 1-phosphate synthase (MI 1-PS) present in clone p5dm1-1ps
 CC (ATCC 97970). The clone was isolated from a cDNA library of soybean line
 CC LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
 CC thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,
 CC raffinose and stachyose. A mutant MI 1-PS nucleic acid (see AAV62443) has
 CC been identified in soybean line LR13, a mutagenised line of low raffinose
 CC saccharide phenotype. Sequencing revealed a single base change mutation
 CC (G to T at base 1241) in the LR13 sequence. The mutation results in a
 CC seed phenotype of very low raffinose saccharide sugars, very high sucrose
 CC and low phytic acid. The nucleic acid is used to alter the raffinose
 CC saccharide, sucrose, phytic acid and inorganic phosphate content of
 CC soybean seeds, leading to useful soybean products, e.g. a seed phytic
 CC acid content of less than 17 ug/g, a seed content of raffinose and
 CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
 CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX Sequence 1782 BP; 516 A; 371 C; 411 G; 484 T; 0 U; 0 Other;
 SQ
 Query Match 9.5%; Score 146; DB 2; Length 1782;
 Best Local Similarity 100.0%; Pred. NO. 1e-60;
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 937 AAAATGAATCTGTGTTGATT 962
DB 990 AAAATGAATCTGTGTTGATT 1015

DB 95 ATCCTATAGAGCTGTGAACATCC 120

Search completed: June 8, 2005, 07:27:00
Job time : 843.149 secs

RESULT 15

ADP93370
ID ADP93370 standard; cDNA; 377 BP.

AC ADP93370;

DT 09-SEP-2004 (first entry)

DE Cotton expressed sequence tag, EST, #2381.

KM Cotton; see EST; expressed sequence tag; plant; plant protection;
KM plant improvement; marker-assisted breeding.

OS Gossypium hirsutum; variety Nucleon33B.

XX US2004123338-A1.

PN 24-JUN-2004.

XX 08-DEC-2000; 2000US-00732627.

XX 10-DEC-1999; 99US-0170255P.

XX (FINC/) FINCHER K L.

XX PA Finch

XX PI Finch

XX DR WPI; 2004-479807/45.

PT New substantially purified nucleic acid molecule that encodes a cotton
PT protein or its fragment, useful as molecular tool for the targeting and
PT isolation of novel genes for plant protection and improvement.

PS Claim 1, SEQ ID NO 2381; 30bp; English.

CC The invention relates to a substantially purified nucleic acid molecule
CC that encodes a cotton protein or its fragment comprising an EST
CC (expressed sequence tag) appearing as ADP90990-ADP95919. Also included
CC are a substantially purified cotton protein or its fragment encoded by a
CC nucleic acid molecule above and a transformed plant (having a nucleic
CC acid molecule which comprises: an exogenous promoter region which
CC functions in a plant cell to cause the production of a mRNA molecule; a
CC structural nucleic acid molecule comprising one of the ESTs or their
CC complements; a 3' non-translated sequence that functions in the plant
CC cell to cause termination of transcription and addition of polyadenylated
CC ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
CC molecular tool for the targeting and isolation of novel genes for plant
CC protection and improvement. The ESTs are useful for developing new
CC strategies for understanding critical plant developmental and metabolic
CC pathways, for isolating genes and promoters, for identifying and mapping
CC the genes involved in developmental and metabolic pathways, and for
CC determining gene function. The cotton nucleic acid molecules are useful
CC as molecular tags to isolate genetic regions, isolate genes, map genes,
CC and determine gene function. The nucleic acid molecules are useful for
CC determining if genes are members of a particular gene family and for use
CC in marker-assisted breeding programs. The present sequence is one of the
CC 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed
CC in the specification but are available in electronic format from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocID=20040123338.

XX Sequence 377 BP; 100 A; 87 C; 90 G; 100 T; 0 U; 0 Other;

Query Match 1.7%; Score 26; DB 12; Length 377;

Best Local Similarity 100.0%; Pred. No. 0.064; Mismatches 0; Indels 0; Gaps 0;

QY 1114 ATCCTATAGAGCTGTGAACATCC 1139

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:49:46 ; Search time 5221.88 Seconds
(without alignments)
11174.613 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgctcatcagagattttaa.....tgattctcagatcaagta 1533

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database : EST:
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	551	35.9	552	5	BM955039 sam76c03.
2	544	35.5	754	4	B1968101 GM830004A
3	543	35.4	594	4	BM523521 sam85a05.
4	540	35.2	540	5	BM887128 sam34e09.
5	517	33.7	561	1	AI442485 aa27f05.Y
6	503	32.4	525	2	BE191344 en71c11.Y
7	496	32.4	581	2	BM085458 ba337a06.
8	479	31.2	782	4	AW348136 GM210001A
9	466	30.4	585	4	BG239011 ba057d09.
10	462	30.1	618	2	AW279066 sg07d02.Y
11	444	29.0	449	2	AM203517 ef35p05.Y
12	419	27.3	632	6	CA938527 ba337a06.
13	413	26.3	516	4	BG652636 ba069g02.
14	374	24.6	505	2	BE660322 1274 Gmax
15	374	24.4	588	5	BUS49101 GM880018B
16	348	22.7	420	4	BM092617 ba116d10.
17	304	19.8	304	4	BF425513 su56f04.Y
18	287	19.4	493	5	BO081344 sam23b12.
19	282	19.0	634	2	BE331050 ba093e10.Y
20	278	18.1	457	2	AW310369 ef35p05.X
21	277	18.1	411	2	AM099866 ba117e12.Y
22	277	18.1	456	2	AM620996 ef35p05.Y
23	274	17.9	274	2	AM153049 ba337a06.Y
24	264	17.2	467	1	AI442850 aa27f05.X

25	247	16.1	670	4	BG044525 aa29e07.
26	232	15.1	539	2	AM184796 ba069g02.Y
27	214	14.0	526	4	B1894109 ba160c05.
28	209	13.6	226	1	A1960904 bc92c07.Y
29	187	12.2	295	4	B1944254 aa95f07.Y
30	186	12.1	187	2	BF596401 su71g11.Y
31	185	12.1	388	2	AM432728 ba093e10.Y
32	179	11.7	475	6	CB063430 ba093e10.Y
33	152	9.9	152	2	BE596037 ba069g02.Y
34	149	9.7	172	2	BE609610 ba069g02.Y
35	146	9.5	308	2	AW705757 ba069g02.Y
36	146	9.5	443	5	BO612082 ba069g02.Y
37	146	9.5	533	4	BM523576 ba069g02.Y
38	146	9.5	661	2	BE191464 en71c11.Y
39	146	9.5	813	7	CK768601 ba069g02.Y
40	135	8.8	290	2	AW460108 ba110d02.Y
41	135	8.8	316	2	AM398011 sg71d06.Y
42	135	8.8	396	2	AM472088 ba119g09.Y
43	135	8.8	619	2	BE311363 ba069g02.Y
44	135	8.8	686	2	AM348857 ba069g02.Y
45	132	8.6	271	2	AM568795 ba161d10.Y

ALIGNMENTS

RESULT 1
LOCUS BM955039 552 bp mRNA linear EST 05-JUL-2004
DEFINITION sam76c03.Y1 Gm-cl069 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl069-5118.5' similar to TR:Q9SSV4 Q9SSV4
MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ; mRNA sequence.

ACCESSION BM955039
VERSION BM955039.1 GI:19453629
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE
AUTHORS 1 (bases 1 to 552)
Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,V., Khanna,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Treistling,B., Allen,M., Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project

JOURNAL
COMMENT Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence etop: 420.
Location/Qualifiers

FEATURES

source
1..552
/organism="Glycine max"
/mol_type="mRNA"
/cultiar="Williams"
/db_xref="Lacox:3847"
/clone="SOYBEAN CLONE ID: Gm-cl069-5118"
/tissue_type="Degenerating cotyledons, 9-10 day old etiolated seedling"
/lab_host="DH10B"

QY 1110 TGGCATCTCTATAGACCTGTGTAACATCCGACATGTTGTTATTATTAAGTATGCC 1169
DB 582 TGGCATCTCTATAGACCTGTGTAACATCCGACATGTTGTTATTATTAAGTATGCC 523
QY 1170 TTACGTAGGGGATAGCAAGAGAGCCATGATGATCACTTCAGAGATATTCATGGGTGG 1229
DB 522 TTACGTAGGGGATAGCAAGAGAGCCATGATGATCACTTCAGAGATATTCATGGGTGG 463
QY 1230 AAAAGAACACATGTTTTCAGCAACATGTGAGATTCCTCTTTAGCTGCTCTATTAT 1289
DB 462 AAAAGAACACATGTTTTCAGCAACATGTGAGATTCCTCTTTAGCTGCTCTATTAT 403
QY 1290 CTTCGATCTGCTCTCTGCTGATGACATTAAGATTCAGTTTAAAGCTGAATGA 1349
DB 402 CTTCGATCTGCTCTCTGCTGATGACATTAAGATTCAGTTTAAAGCTGAATGA 343
QY 1350 GGGAAATTCATCTCATTCACCCAGTTGCTACATTCAGTCTATCTGACCAAGGCTCC 1409
DB 342 GGGAAATTCATCTCATTCACCCAGTTGCTACATTCAGTCTATCTGACCAAGGCTCC 283
QY 1410 TCTGTTCCACCGGGTACACACAGTGTGAATGATTTGTCAAGACCGTGCATGCTGA 1469
DB 282 TCTGTTCCACCGGGTACACACAGTGTGAATGATTTGTCAAGACCGTGCATGCTGA 223
QY 1470 AACCATATGAGGGCTTGTGTTGATTGGCCGAGAAATTAACATGATTCGAGTACA 1529
DB 222 AACCATATGAGGGCTTGTGTTGATTGGCCGAGAAATTAACATGATTCGAGTACA 163
QY 1530 GTGA 1533
DB 162 GTGA 159

RESULT 3
LOCUS BM523521 594 bp mRNA linear EST 06-JUL-2004
DEFINITION sam5a05.y2 Gm-cl087 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl087-777 5' similar to TR:09SSV4 09SSV4 MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ;, mRNA sequence.

ACCESSION BM523521
VERSION BM523521.1 GI:18724004
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 594)
AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Paape,D., Harvey,N., Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McGinn,R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave, Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence scop: 434.

FEATURES
source 1..594
Location/Qualifiers
/organism="Glycine max"

/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl087-777"
/issue_type="Roots"
/lab_host="DH10B"
/clone_1b="Gm-cl087"
/note="vector: pBluescript II SK+; Site 1: EcoRI, Site 2: XhoI; The mRNA was prepared using poly(A) tail extract mRNA system from PROMEGA. The cDNA was prepared using the STRATAGENE kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site (5'GAGAGAGAGAGAGAGATGATCTGAGTTTCTTTTCTTTT). EcoRI adapters (5'OH-AATTCGACGACG and 3'GCCGCTCTC) were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction sites of the pBluescript vector. The vector was previously dephosphorylated. The ligated cDNA fragments were transformed into DH10B by electroporation. Library is in LB medium with ampicillin and glycerol 8%. Average insert size: 800 bp."

Query Match 35.4%; Score 543; DB 4; Length 594;
Best Local Similarity 99.8%; Pred. No. 5.7e-300;
Matches 593; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 847 CCTCAGAACACTTTTGTACCAAGGCTGATGATCTTGCATCGCAGAACACTTTGATT 906
DB 1 CCTCAGAACACTTTTGTACCAAGGCTGATGATCTTGCATCGCAGAACACTTTGATT 60

QY 907 GGTGAGATGACTCAAGAGTGTGACAGACCAAAATGAAATCTGTGTGATTTTCTT 966
DB 61 GGTGAGATGACTCAAGAGTGTGACAGACCAAAATGAAATCTGTGTGATTTTCTT 120

QY 967 GTGGGGGCGGTATCAAGCAACATCTATGATTACCAACATCTGGGAAACATGAT 1026
DB 121 GTGGGGGCGGTATCAAGCAACATCTATGATTACCAACATCTGGGAAACATGAT 180

QY 1027 GGTATGATCTCTCGCTCCACAACTTCGCTCAAGGAATCTCCAAAGCAACGTT 1086
DB 181 GGTATGATCTCTCGCTCCACAACTTCGCTCAAGGAATCTCCAAAGCAACGTT 240

QY 1087 GTTGACGATATGTCAACAGCAATCCATCTTATGAGCTGTGACATCCGACAT 1146
DB 241 GTTGACGATATGTCAACAGCAATCCATCTTATGAGCTGTGACATCCGACAT 300

QY 1147 GTTGTGTTATTAAGTATGCTTACGTAAGGGGATAGCAAGGCAATGATGATC 1206
DB 301 GTTGTGTTATTAAGTATGCTTACGTAAGGGGATAGCAAGGCAATGATGATC 360

QY 1207 ACTTCAGAGATATTTGATGAGTGTGAAAGAACACATTTGTTTTCACAAACATGAGAT 1266
DB 361 ACTTCAGAGATATTTGATGAGTGTGAAAGAACACATTTGTTTTCACAAACATGAGAT 420

QY 1267 TCCCTTTTAACTGCTCTTATATCTTGAATGATGCTTCTTGTGAGCTGAGCACTAGA 1326
DB 421 TCCCTTTTAACTGCTCTTATATCTTGAATGATGCTTCTTGTGAGCTGAGCACTAGA 480

QY 1327 ATCCAGTTTAAAGCGAAAGAGGAAATTCACATTCACACCGCTGCTACCAT 1386
DB 481 ATCCAGTTTAAAGCGAAAGAGGAAATTCACATTCACACCGCTGCTACCAT 540

QY 1387 CTACACTATCTGACCAAGGCTCTCTGTTTCCACCGGTTACACAGTGTGAT 1440
DB 541 CTACACTATCTGACCAAGGCTCTCTGTTTCCACCGGTTACACAGTGTGAT 594

RESULT 4
LOCUS BM887128 540 bp mRNA linear EST 06-JUL-2004
DEFINITION sam5a05.y1 Gm-cl068 Glycine max cDNA clone SOYBEAN CLONE ID:


```

ACCESSION   Gm-c1068-6738 5' similar to TR:Q9SSV4 Q9SSV4
VERSION      MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ;, mRNA sequence.
KEYWORDS    BM887128
SOURCE       EST.
ORGANISM     Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE    1 (bases 1 to 540)
AUTHORS      Shoemaker,R., Kelm,P., Vodkin,L., Erpelidng,J., Coryell,V.,
             Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
             Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
             Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
             Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
             McCann,R., Waterston,R. and Wilson,R.
TITLE        Public Soybean EST Project
JOURNAL      Unpublished (1999)
COMMENT      Contact: Shoemaker R./Public Soybean EST Project
             Public Soybean EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             When it has been determined, an EST from the other end of this
             clone is listed in the 'Other ESTs on clone' field. This clone is
             available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
             57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
             Seq primer: -40RP from Gibco
             High quality sequence stop: 421.

FEATURES
    source
        1..540
            /organism="Glycine max"
            /mol_type="mRNA"
            /cultivar="Williams 82"
            /db_xref="taxon:3847"
            /clone="SOYBEAN CLONE ID: Gm-c1068-6738"
            /tissue_type="leaf; drought stressed, 1 month old plants,
            greenhouse grown"
            /lab_host="DH10B"
            /clone_lib="Gm-c1068"
            /note="Vector: pBluescript II SK+; Site 1: EcoRI, Site 2:
            XhoI. The cDNA library was constructed from mRNA isolated
            from drought stressed leaf tissue of the cultivar Williams
            82. The month old greenhouse grown plants were deprived of
            water for 3 days prior to harvesting the stressed leaf
            tissue. Complementary DNA was synthesized from mRNA using
            a primer consisting of a poly(dT) sequence with a XhoI
            restriction site. EcoRI adapters were ligated to the
            blunt-ended cDNA fragments followed by XhoI digestion. The
            cDNA fragments were directionally cloned into the
            EcoRI-XhoI restriction site of the pBluescript vector. The
            ligated cDNA fragments were transformed into DH10B host
            cells (GibcoBRL). This library was constructed in the
            laboratory of Dr. Randy Shoemaker."

ORIGIN
Query Match      35.2%; Score 540; DB 5; Length 540;
Best Local Similarity 100.0%; Pred. No. 3e-298;
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 598 GGCACAAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGCGCTTTAAGAGAGCC 657
      |||||||
DB 1 GGCACAAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGCGCTTTAAGAGAGCC 60
      |||||||
QY 658 ACCAAAGTGGACAAAGGTGTGTGCTGTGACAGCCAAACAGAGAGGATATGCAATTGG 717
      |||||||
DB 61 ACCAAAGTGGACAAAGGTGTGTGCTGTGACAGCCAAACAGAGGATATGCAATTGG 120
      |||||||
QY 718 GTTGTAGGCTTAATGACACATGAGAAATCTCTTGCTGTGTGACAGAAATGAGGCT 777
      |||||||

ACCESSION   A1442485
VERSION      A1442485.1 GI:4296509
KEYWORDS
SOURCE       Glycine max (soybean)
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
             Glycine.
REFERENCE    1 (bases 1 to 561)
AUTHORS      Shoemaker,R., Kelm,P., Vodkin,L., Erpelidng,J., Coryell,V.,
             Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
             Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
             Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
             Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
             McCann,R., Waterston,R. and Wilson,R.
TITLE        Public Soybean EST Project
JOURNAL      Unpublished (1999)
COMMENT      Other_ESTs: AM348136 corresponding to Gm-r1021-68 (3')
             Contact: Shoemaker R./Public Soybean EST Project
             Public Soybean EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             When it has been determined, an EST from the other end of this
             clone is listed in the 'Other ESTs on clone' field. Possible
             reversed clone: similarity on wrong strand This clone is available
             through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA
             (phone: 800 423 4163; email: info@biogeneticservices.com)
             Insert Length: 1557 Std Error: 0.00
             Seq primer: -40RP from Gibco
             High quality sequence stop: 390
             POLYA=No.

FEATURES
    source
        1..561
            /organism="Glycine max"
            /mol_type="mRNA"

```

/cultivar="Williams"
 /db_xref="taxon:3841"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl004-538"
 /tissue_type="root"
 /lab_host="XL10-Gold"
 /clone_id="Gm-cl004"
 /note="Vector: plasmidscript II XR, Site 1: EcoRI; Site 2:
 XhoI; Root cDNA. The mRNA was isolated from entire roots
 of 8 day old 'Williams' seedlings which were propagated on
 paper towels with distilled water. Stratsagene's cDNA
 Synthesis Kit (catalog #200401) was used to synthesize the
 cDNA. First- strand synthesis was performed with 5-methyl-
 dCTP, hence the ligated cDNA is hemimethylated.
 Stratsagene's first-strand synthesis primer was used
 (GAGACGAGACGAGACGAGACGACTGCTCGAG(T)-18). After
 second-strand synthesis, the cDNA ends were 'polished'
 with clone pT4 DNA polymerase, ligated to EcoRI adaptors,
 and phosphorylated. The XhoI site within the first-strand
 synthesis primer was restricted by digestion with XhoI;
 all XhoI sites in the cDNA would be protected by their
 hemimethylated status. The cDNA constructs were
 size-fractionated with a 500pg cutoff, using GibcoBRL Life
 Technologies' cDNA Size Fractionation column. The column
 eluent was then ligated into Stratsagene's plasmidscript II
 XR predigested vector (plasmidscript II SK(+), that had been
 digested with EcoRI and XhoI, and phosphorylated). Both
 the white and blue colonies appear to contain recombinant
 plasmids with cDNA inserts. Blue colonies 9n-15 have been
 sequenced, and possess putative cDNA inserts. This library
 was constructed by Dr. Paul Keim & Virginia H. Coryell,
 Department of Biology, Box5660, Northern Arizona
 University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr.
 Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax:
 520-523-7500, email: paul.keim@nau.edu,
 virginia.coryell@nau.edu"

ORIGIN

Query Match	33.7%	Score 517;	DB 1;	Length 561;
Best Local Similarity	100.0%;	Pred. No. 5,1e-285;		
Matches 517;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	341	GATCCTTCAGAGAGAGAAAATCTATGCCCCCATTCAGAGAGTCGTCTCCAAATGTTAATC	400	
Db	5	GATCCTTCAGAGAGAGAAAATCTATGCCCCCATTCAGAGAGTCGTCTCCAAATGTTAATC	64	
QY	401	CTGACGACATTTGTGTTGGGGGATGGAATATCAGCAACATGAACCTGGCTGAATGCAATGG	460	
Db	65	CTGACGACATTTGTGTTGGGGGATGGAATATCAGCAACATGAACCTGGCTGAATGCAATGG	124	
QY	461	CCAGGGCAAAAGTGTTTGACATCCATTTGCAAGAGACAGTTGAGGCTTACATGAAATCCA	520	
Db	125	CCAGGGCAAAAGTGTTTGACATCCATTTGCAAGAGACAGTTGAGGCTTACATGAAATCCA	184	
QY	521	TGTTTCCACTCCCGGAATCTACGACCGGGAATTCATATGCTGCGCAACCAAGAGAGCGGTG	580	
Db	185	TGTTTCCACTCCCGGAATCTACGACCGGGAATTCATATGCTGCGCAACCAAGAGAGCGGTG	244	
QY	581	CCAACAACGTATTTAAGGGCACAAAGCAAGACAAATTACAGCAATATCTCAAAGACATCA	640	
Db	245	CCAACAACGTATTTAAGGGCACAAAGCAAGACAAATTACAGCAATATCTCAAAGACATCA	304	
QY	641	AGGCGTTTAAAGAAAGCCACCAAAAGTGACAAAGTGTGTCCTGTGACTGCCAACAG	700	
Db	305	AGGCGTTTAAAGAAAGCCACCAAAAGTGACAAAGTGTGTCCTGTGACTGCCAACAG	364	
QY	701	AGAGGTATAGCAATTTGGTGTGAGGCTTAAATGACACATGAGAAATCTTTGGCTGCTG	760	
Db	365	AGAGGTATAGCAATTTGGTGTGAGGCTTAAATGACACATGAGAAATCTTTGGCTGCTG	424	
QY	761	TGGAAGAAATGAGGCTGAGATTTCTCCTTCCACTTGTATGSCATTTGCTGTGTATGG	820	
Db	425	TGGAAGAAATGAGGCTGAGATTTCTCCTTCCACTTGTATGSCATTTGCTGTGTATGG	484	

DB	Accession	Version	Keywords	Organism	Source	Reference	Authors	Title	Journal	Comment
821	AAATGTTCCTTTCATTAAAGGAAGCCCTAGAACAC	857								
485	AAATGTTCCTTTCATTAAAGGAAGCCCTAGAACAC	521								
RESULT 6	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	SOURCE	REFERENCE	AUTHORS	TITLE
BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344
525 bp	525 bp	525 bp	525 bp	525 bp	525 bp	525 bp	525 bp	525 bp	525 bp	525 bp
linear	linear	linear	linear	linear	linear	linear	linear	linear	linear	linear
EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004	EST 14-JUL-2004
GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:	GENOME SYSTEMS CLONE ID:
Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196	Gm-c1038-45 5' similar to TR:065196
MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE	MYO-INOSITOL 1-PHOSPHATE
SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.	SYNTHASE ; mRNA sequence.
BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344	BE191344
GI:8670237	GI:8670237	GI:8670237	GI:8670237	GI:8670237	GI:8670237	GI:8670237	GI:8670237	GI:8670237	GI:8670237	GI:8670237
EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.
Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)	Glycine max (soybean)
Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.	Eukaryota ; Viridiplantae ; Streptophyta ; Embryophyta ; Tracheophyta ; Spermatophyta ; Magnoliophyta ; eudicotyledons ; core eudicots ; rosids ; eurosids I ; Fabales ; Fabaceae ; Papilionoideae ; Phaseoleae ; Glycine.
(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)	(bases 1 to 525)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, U., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Stepcie, W., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R., and Wilson, R.	Shoemaker, R., Keim, P., Vodkin, L., Erpelting, U., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Stepcie, W., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R., and Wilson, R.	Shoemaker, R., Keim, P., Vodkin, L., Erpelting, U., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Stepcie, W., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R., and Wilson, R.	Shoemaker, R., Keim, P., Vodkin, L., Erpelting, U., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Stepcie, W., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R., and Wilson, R.	Shoemaker, R., Keim, P., Vodkin, L., Erpelting, U., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Stepcie, W., Theising						

FEATURES	COMMENT
source	<p>TITLE Public Soybean EST Project</p> <p>JOURNAL Unpublished (1999)</p> <p>Contact: Shoemaker R/Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: east@wustl.edu</p> <p>When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4153; email: info@biogeneticservices.com)</p> <p>Insert Length: 700 Std Error: 0.00</p> <p>High quality sequence stop: 490.</p> <p>Location/Qualifiers 1. .525</p>

```

/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl038-45"
/tissue_type="Senescing leaves, greenhouse grown"
/dev_stage="mature plants"
/lab_host="DH10B"
/clone_1lb="Gm-cl038"
/name="Vector: pT773Pac (pT773, Pharmacia); Site_1: EcoRI; Site_2: HindIII; This cDNA library was constructed from mRNA isolated from the senescing leaf tissue of mature greenhouse grown plants of the cultivar Williams 82. Complementary DNA was synthesized from mRNA using a 3', anchored poly(dT) primer. EcoRI adapters were ligated to the blunt ended cDNA fragments followed by digestion with EcoRI and HindIII. The cDNA fragments were directionally cloned into the EcoRI-HindIII restriction site of the pT773-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker and Dr. John Eppelring."

```

ORIGIN	
Query Match	32.8%; Score 503; DB 2; Length 525;
Best Local Similarity	100.0%; Pred. No. 5 8e-277;
Matches 503; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
0y	1031 TGAATCTCTCGGCTCACAAACCTTCGGTTCAGAGAAATCTCCAGACGACCTTGTTG 1090

```
Db      4 TGAATCTCGGCTCCACAACCTTCCGCTCCAGGAATCTCCAAGACAAGCTTGTG 63
Qy      1091 AGGATATGCTCAACAGCAATGCATCCTATAGAGCTGTGTAACAATCCGACCATTTG 1150
Db      64 AGGATATGCTCAACAGCAATGCATCCTATAGAGCTGTGTAACAATCCGACCATTTG 123
Qy      1151 TTGTTATTAAGTATGCTTACCTAGTAGGGATAGCAAGACCATGATGATCACTT 1210
Db      124 TTGTTATTAAGTATGCTTACCTAGTAGGGATAGCAAGACCATGATGATCACTT 183
Qy      1211 CAGAGATATTCATGGGTGGAAGAACCAATTTTGGACAACAACATGTAGATTTCC 1270
Db      184 CAGAGATATTCATGGGTGGAAGAACCAATTTTGGACAACAACATGTAGATTTCC 243
Qy      1271 TTTTTCGCTGCTCTATTAATCTTGAATGCTTCTTCTGATGATGATCACTTAAGTCC 1330
Db      244 TTTTTCGCTGCTCTATTAATCTTGAATGCTTCTTCTGATGATGATCACTTAAGTCC 303
Qy      1331 AGTTTAAAGCTGAAATGAGGAAATTCACATTCACCAAGTGTACCATTTCTCA 1390
Db      304 AGTTTAAAGCTGAAATGAGGAAATTCACATTCACCAAGTGTACCATTTCTCA 363
Qy      1391 GCTATCTGACCAAGCTCTCTGTGTTCCACCGGATACACCAAGTGTGAATGATGTCAA 1450
Db      364 GCTATCTGACCAAGCTCTCTGTGTTCCACCGGATACACCAAGTGTGAATGATGTCAA 423
Qy      1451 AACAGGCTGCATTCGTGGAAGAAATATAGAGGCTTGTGATGATGGCCGACGAATA 1510
Db      424 AACAGGCTGCATTCGTGGAAGAAATATAGAGGCTTGTGATGATGGCCGACGAATA 483
Qy      1511 ACATGATTTCTCGATGATCAAGTGA 1533
Db      484 ACATGATTTCTCGATGATCAAGTGA 506
```

```
RESULT 7
LOCUS   BM085458
DEFINITION
541 bp mRNA linear EST 23-JUL-2004
sa372a06.y1 Gm-c1066 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1066-4644.5, similar to TR:095SV4 Q9SSV4
MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ; mRNA sequence.
BM085458
BM085458.1 GI:16996086
EST.
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 541)
Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Stephens,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibson,M., Peape,D., Harvey,N.,
Schurk,R., Riter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watsn.wustl.edu
When it was determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
High quality sequence stop: 426.
Location/Qualifiers
```

```
source
1. 541
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1066-4644"
/tissue_type="Leaf and shoot tip, salt stressed, 2 week
old seedling"
/lab_host="DH10B"
/clone_id="Gm-c1066"
/note="Vector: pBluescript II SK+, Site 1: EcoRI, Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from unexpanded leaves and the shoot tips of 2 week old
seedling from the cultivar Williams. The 2 week old
seedlings were salt stressed in a solution of 500mM NaCl
for 3 days prior to harvesting. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."
```

Query Match 32.4%; Score 496; DB 4; Length 541;

Best Local Similarity 100.0%; Pred. No. 6.2e-273; Mismatches 0; Indels 0; Gaps 0;

Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1038 CTCGGCTCCAGAAACCTTCGCTCCAGAAATCTCAAGACAACTGTTGAGATAT 1097

Db 1 CTCGGCTCCAGAAACCTTCGCTCCAGAAATCTCAAGACAACTGTTGAGATAT 60

Qy 1098 GGTCAACAGCAATGCTATCTTATGAGCTGTGTGAATATCCGACATGTTGTTAT 1157

Db 61 GGTCAACAGCAATGCTATCTTATGAGCTGTGTGAATATCCGACATGTTGTTAT 120

Qy 1158 TAAAGTATGCTTACGTAAGGGGATAGCAAGAGCCATGTAGTACATTCAGAGAT 1217

Db 121 TAAAGTATGCTTACGTAAGGGGATAGCAAGAGCCATGTAGTACATTCAGAGAT 180

Qy 1218 ATTCAATGAGTGAAGAACACATGTTTTCACAAACATGTAGATTCCTTTTAC 1277

Db 181 ATTCAATGAGTGAAGAACACATGTTTTCACAAACATGTAGATTCCTTTTAC 240

Qy 1278 TGCTCTTATTTATCTTGAATGCTTCTTCTTGTGAGTGAAGTACATTAATTC 1337

Db 241 TGCTCTTATTTATCTTGAATGCTTCTTCTTGTGAGTGAAGTACATTAATTC 300

Qy 1338 AGCTGAATAAGAGGAAATTCATCTCATTCACACCACTTGTGATTCAGATATCT 1397

Db 301 AGCTGAATAAGAGGAAATTCATCTCATTCACACCACTTGTGATTCAGATATCT 360

Qy 1398 GACCAAGCTCTCTGTTTCCACCGGTATACCAAGTGTGATGATGTCAAAAGCAG 1457

Db 361 GACCAAGCTCTCTGTTTCCACCGGTATACCAAGTGTGATGATGTCAAAAGCAG 420

Qy 1458 TGCAATGCTGAAACATTAAGAGGCTTGTGTGATGATGAGCCCAAGATTAATAT 1517

Db 421 TGCAATGCTGAAACATTAAGAGGCTTGTGTGATGATGAGCCCAAGATTAATAT 480

Qy 1518 TCTCGAGTACAAATGA 1533

Db 481 TCTCGAGTACAAATGA 496

```
RESULT 8
LOCUS   AM348136/c
DEFINITION
782 bp mRNA linear EST 04-OCT-2000
AM348136 Gm210001A21B10 Gm-r1021 Glycine max cDNA clone Gm-r1021-68 3', mRNA
sequence.
```

ACCESSION AM348136
 VERSION AM348136.1 GI:6845846
 KEYWORDS EST
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max
 BUKARYOTA; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 1 (bases 1 to 782)
 Vodka, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V., Erpelting, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L., and Lewin, H.
 A Functional Genomics Program for Soybean (NSF 9872565)
 Unpublished (1999)
 Other ESTs: A1442485
 Contact: Vodka, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vodka@uiuc.edu
 This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genome systems.com web site: www.genomesystems.com
 Seq primer: 5'-TTTTTTTTTTTTTTTTTTT (A/C/G) -3'.
 Location/Qualifiers
 1..782
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Williams"
 /db_xref="taxon:3847"
 /clone="Gm-r1021-68"
 /tissue_type="root"
 /lab_host="X10-Gold"
 /clone_lib="Gm-r1021"
 /note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; Library Gm-r1021 is a sequence-driven, rechecked set of the original library Gm-cl004 which was prepared from root cDNA. The mRNA was isolated from entire roots of 8 day old 'Williams' seedlings which were propagated on paper towels with distilled water. Stragene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. The Gm-cl004 library was constructed by Dr. Paul Keim & Virginia H. Coryell, Department of Biology, Box5640, Northern Arizona University, Flagstaff, AZ 86011, email: paul.keim@uau.edu, virginia.coryell@uau.edu. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Computational Biology Centers, University of Minnesota.
 http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html
 . Retzel, E. was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois.
 http://www.life.uiuc.edu/biotech/keck.html."

ORIGIN
 Query Match 31.2%; Score 479; DB 2; Length 782;
 Best Local Similarity 100.0%; Pred. No. 4e-263;
 Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TCCGCTCCAGGAATCTCCAGAGCAAGCTTGAACATATGTCACACAGAAAGCA 1114
 DB 642 TCCGCTCCAGGAATCTCCAGAGCAAGCTTGAACATATGTCACACAGAAAGCA 583
 QY 1115 TCCCTATAGAGCTGTGAACATCCGACCATGTTGTTATTAAGTATGTCCTTACG 1174
 DB 582 TCCCTATAGAGCTGTGAACATCCGACCATGTTGTTATTAAGTATGTCCTTACG 523

QY 1175 TAGGGATATGCAAGAGGCAATGATGATACCTTCAGAGATATTCATGCGTGAAGA 1234
 DB 522 TAGGGATATGCAAGAGGCAATGATGATACCTTCAGAGATATTCATGCGTGAAGA 463
 QY 1235 AACACATGTTTGTGCAACACATGAGAGATTCCTTTACTGCTGCTATATCTTGG 1294
 DB 462 AACACATGTTTGTGCAACACATGAGAGATTCCTTTACTGCTGCTATATCTTGG 403
 QY 1295 ACTTGTCCTTCTTGTGAGTACAGCACTGATGATGATGATGATGATGATGATG 1354
 DB 402 ACTTGTCCTTCTTGTGAGTACAGCACTGATGATGATGATGATGATGATGATG 343
 QY 1355 AATTCATCTATTCACACCATGTTGCTACATTCATCTGATTCGACCAAGCTCTG 1414
 DB 342 AATTCATCTATTCACACCATGTTGCTACATTCATCTGATTCGACCAAGCTCTG 283
 QY 1415 TTCCACCGGCTACACAGTGTGAATGATGATGATGATGATGATGATGATGATG 1474
 DB 282 TTCCACCGGCTACACAGTGTGAATGATGATGATGATGATGATGATGATGATG 223
 QY 1475 TAATGAGGCTTGTGTTGATGCGCCGACAGATTAATGATTCGATGCAAGTGA 1533
 DB 222 TAATGAGGCTTGTGTTGATGCGCCGACAGATTAATGATTCGATGCAAGTGA 164
 RESULT 9
 BG239011
 LOCUS
 DEFINITION
 585 bp mRNA linear EST 23-JUL-2004
 bab57d09.y1 Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-cl043-3258 5' similar to TF:Q5SSV4 Q5SSV4
 MYO-INOSITOL-1-PHOSPHATE SYNTHASE.; mRNA sequence.
 BG239011
 BG239011.1 GI:12774084
 EST.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Glycine max (soybean)
 BUKARYOTA; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 1 (bases 1 to 585)
 Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Treising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurr, R., Ritter, B., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
 High quality sequence stop: 436.
 Location/Qualifiers
 1..585
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Williams"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl043-3258"
 /tissue_type="Hypocotyl and Plumule, germinating seeds"
 /lab_host="DH10B"
 /clone_lib="Gm-cl043"
 /note="Vector: pT73Pac (Pharmacia); Site_1: EcoRI;

Site 2: NotI; This cDNA library was constructed from mRNA isolated from hypocotyl and plumule tissues of seeds generated for three days of the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the pT73-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

ORIGIN

Query Match 30.4%; Score 466; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 893 GGAGACCTTTGATGTTGAGATGACCTTCAAGAGTGTGACCAAAATGAAATCTGTGT 952
DB 14 GGAGACCTTTGATGTTGAGATGACCTTCAAGAGTGTGACCAAAATGAAATCTGTGT 73
QY 953 TGGTGGATTTTCTTGTGGGGGCTGGTATCAACCAACATCTTATGTTTCAACCATC 1012
DB 74 TGGTGGATTTTCTTGTGGGGGCTGGTATCAACCAACATCTTATGTTTCAACCATC 133
QY 1013 TGGGAACAATGATGATGATGATCTCTGGGCTCCAAACCTTCGGCTCCAGAAATCT 1072
DB 134 TGGGAACAATGATGATGATGATCTCTGGGCTCCAAACCTTCGGCTCCAGAAATCT 193
QY 1073 CCAAGAGCAACGTTGTCAGATGTCAGACAGCAATGCCATCTTATGAGCCTGTG 1132
DB 194 CCAAGAGCAACGTTGTCAGATGTCAGACAGCAATGCCATCTTATGAGCCTGTG 253
QY 1133 AACATCCGACCATGTTGTTTATTAAGTATGCTTACGTAGGGATACAGAGAG 1192
DB 254 AACATCCGACCATGTTGTTTATTAAGTATGCTTACGTAGGGATACAGAGAG 313
QY 1193 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1252
DB 314 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 373
QY 1253 ACACATGTTGAGGATTCCTTTTAACTGCTCTTATTAATCTTGAATCTTCTTGTG 1312
DB 374 ACACATGTTGAGGATTCCTTTTAACTGCTCTTATTAATCTTGAATCTTCTTGTG 433
QY 1313 ACCTGACCTGATGATCCAGTTTAAAGCTGAAAATGAGGAAAATT 1358
DB 434 ACCTGACCTGATGATCCAGTTTAAAGCTGAAAATGAGGAAAATT 479

RESULT 10

AM279066

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST.

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 618)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Riter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

AUTHORS

REFERENCE

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
CONTACT: Shoemaker R/Public Soybean EST Project
COMMENT Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1547 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 441.
Location/Qualifiers

FEATURES

source

1..618
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-4924"
/csequence_type="immature seed coats of greenhouse grown plants"
/lab_host="DH10B (Gibco BRL)"
/clone_idb="Gm-cl019"
/note="Vector: pSPORT1 (Life Technologies); Site 1: Not I; Site 2: Sal I; This cDNA library was constructed from mRNA isolated from immature seed coats (200-300 mgs) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells (Gibco BRL). This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

ORIGIN

Query Match 30.1%; Score 462; DB 2; Length 618;
Best Local Similarity 99.8%; Pred. No. 2.3e-253;
Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 261 CATTTCATGAGGCTACAAAGACAGATTCACAAAGCCATTAATCTTGGCTCCCTACCCA 320
DB 45 CATTTCATGAGGCTACAAAGACAGATTCACAAAGCCATTAATCTTGGCTCCCTACCCA 104
QY 321 AGCCTCAGCTAATGAGTGTGATCTCTCAGAGGAGAGAAATCTATGCCCCATCAAG 380
DB 105 AGCCTCAGCTAATGAGTGTGATCTCTCAGAGGAGAGAAATCTATGCCCCATCAAG 164
QY 381 TCTGCTTCAATGTTAATCTCTGACGATTTGTTGGGGAGTGGATATCAGCAAT 440
DB 165 TCTGCTTCAATGTTAATCTCTGACGATTTGTTGGGGAGTGGATATCAGCAAT 224
QY 441 GAACCTGCTATGCTATGCTGAGGAGGCAAGGTTTGAATTCGATTTGCGAAGCAT 500
DB 225 GAACCTGCTATGCTATGCTGAGGAGGCAAGGTTTGAATTCGATTTGCGAAGCAT 284
QY 501 GAGGCTTACATGAAATCATGATTCCTCCGCGAATCTACGACCCGGAATTCATGCG 560
DB 285 GAGGCTTACATGAAATCATGATTCCTCCGCGAATCTACGACCCGGAATTCATGCG 344
QY 561 TGCCAACAAGAGAGAGCTGCAACAAGTATTAAGGGCAACAAGCAAGCAAGTCA 620
DB 345 TGCCAACAAGAGAGAGCTGCAACAAGTATTAAGGGCAACAAGCAAGCAAGTCA 404
QY 621 GCAATCATCAAGACATCAAGCGTTTAAAGAAACCAACAAGTGACAAAGTGTGTTGT 680

Db	Accession	Definition	LOCUS	Accession	Version	Keywords	Source	Reference	Authors	Title	Journal	Comment
Db	405	GCAATCATCAAGACATCAAGCCCTTTAAAGAAAGCCACCAAAAGTGACAAAGTGAGTGTGT	449 bp	AM203517	1	GI:5602144	EST.	Glycine max (soybean)	Shoemaker, R.R., Kelm, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kueba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	
Db	681	CCTGTGACCTCCCAACACAGAGGTATAGCAATTTGGTTGAGCCTTAATGACCAT	449 bp	AM203517	1	GI:5602144	EST.	Glycine max (soybean)	Shoemaker, R.R., Kelm, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kueba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	
Db	465	CCTGTGACCTCCCAACACAGAGGTATAGCAATTTGGTTGAGCCTTAATGACCAT	449 bp	AM203517	1	GI:5602144	EST.	Glycine max (soybean)	Shoemaker, R.R., Kelm, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kueba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	
Db	741	GGAGATCTCTTGCGCTGCTGTGGACAGAAATGA	449 bp	AM203517	1	GI:5602144	EST.	Glycine max (soybean)	Shoemaker, R.R., Kelm, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kueba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	
Db	525	GGAGATCTCTTGCGCTGCTGTGGACAGAAATGA	449 bp	AM203517	1	GI:5602144	EST.	Glycine max (soybean)	Shoemaker, R.R., Kelm, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kueba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449
FEATURES	Source	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449	Location/Qualifiers	1..449		

subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI, all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into StrataGene's pBluescript II XR predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Corryell."

ORIGIN	Query Match	29.0%; Score 444; DB 2; Length 449;
	Best Local Similarity 100.0%; Pred. No. 5.1e-243; Indels 0; Gaps 0;	
	Matches 444; Conservative 0; Mismatches 0;	
QY	634 GACATCAAGCGCTTTTAAAGAACCCAAAGTGACAAAGGTGTGTCTCTGTGACCTGCC 693	
Db	6 GACATCAAGCGCTTTTAAAGAACCCAAAGTGACAAAGGTGTGTCTCTGTGACCTGCC 65	
QY	694 AACACAGAGCGGTATAGCAATTTGTGTGAGCCTTATATGACACCATGAGAAATCTCTTG 753	
Db	66 AACACAGAGCGGTATAGCAATTTGTGTGAGCCTTATATACACCATGAGAAATCTCTTG 125	
QY	754 GCTGCTGTGGACAGAAATGAGGCTGAGATTTCTCCACCTTGTAAGCAATTCCTGT 813	
Db	126 GCTGCTGTGGACAGAAATGAGGCTGAGATTTCTCTCCACCTTGTAATGCAATTCCTGT 185	
QY	814 GTGATGGAAATGTCCTTTCATTATAGAACCCCTCAGAACCTTTTGTATACAGGCTG 873	
Db	186 GTGATGGAAATGTCCTTTCATTATAGAACCCCTCAGAACCTTTTGTATACAGGCTG 245	
QY	874 ATTGATCTTCCCATTCGGAGGAACACTTGAATTGTGTGAGATGACTTCAAGAGTGCTAG 933	
Db	246 ATTGATCTTCCCATTCGGAGGAACACTTGAATGTGTGAGATGACTTCAAGAGTGCTAG 305	
QY	934 ACCAAATGAAATCTGTGTGTGATTTTCTGTGGGGCTGGTATCAAGCCAAACATCT 993	
Db	306 ACCAAATGAAATCTGTGTGTGATTTTCTGTGGGGCTGGTATCAAGCCAAACATCT 365	
QY	994 ATATGTTAGTTACAACCATCTGGGAAACAATGATGTATGTAATCTCTGGCTCCCAAAAC 1053	
Db	366 ATATGTTAGTTACAACCATCTGGGAAACAATGATGTATGTAATCTCTGGCTCCCAAAAC 425	
QY	1054 TTCGGCTCCAAAGGAATCTCCAAG 1077	
Db	426 TTCGGCTCCAAAGGAATCTCCAAG 449	

RESULT 12	CA938527	632 bp	mRNA	linear	EST 01-JUL-2004
LOCUS	CA938527				
DEFINITION	CA938527	632 bp	mRNA	linear	EST 01-JUL-2004
VERSION	CA938527				
KEYWORDS	CA938527				
SOURCE	CA938527.1	GI:27427007			
ORGANISM	EST.				
	Glycine max (soybean)				
	Glycine max				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;				
	Glycine.				
REFERENCE	1 (bases 1 to 632)				
AUTHORS	Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Coryell, V.,				
	Khamra, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,				
	Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,				
	Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,				

TITLE
JOURNAL
COMMENT

Schurk, R., Riter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCam, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 447.

FEATURES

source

1. 632
Location/Qualifiers

/organism="Glycine max"
/mol_type="mRNA"
/cultivar="1157"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl023-6165"
/tissue_type="seed coats of greenhouse grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl023"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from seed
coats (100-200mg) of greenhouse grown plants. The
library was prepared using the Life Technologies
pSuperScript cDNA library construction kit. Complementary
DNA was synthesized from mRNA using a poly (dT) sequence
with a Not I restriction site. Sal I linkers adapters
were ligated to the blunt-ended cDNA fragments followed by
Not I digestion. The cDNA fragments were directionally
cloned into the Not I-Sal I restriction site of the
pSPORT1 vector. The ligated cDNA fragments were
transformed into E.coli Electromax DH10B host cells. This
library was constructed by Dr. Lila Vodkin and Dr. Anu
Khanna."

ORIGIN

Query Match 27.3%; Score 419; DB 6; Length 632;
Best Local Similarity 99.8%; Pred. No. 1.3e-228;
Matches 469; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
1 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
111 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 170
61 CAGTCGGTGTACAACTACGAAACCAACCGAACTTGTTCAGAGAACGAAATGGACCTAT 120
171 CAGTCGGTGTACAACTACGAAACCAACCGAACTTGTTCAGAGAACGAAATGGACCTAT 230
121 CAGTGGATTTGCAAAACCAATCCGTCACTCAATTTTAAACCAACCAACCGATTTCCA 180
231 CAGTGGATTTGCAAAACCAATCCGTCACTCAATTTTAAACCAACCAACCGATTTCCA 290
181 AAATTTGGGGGTGATGTTGTGGGTGGGGTGAACAACAGGCTTACCTTCCACCGTGGT 240
291 AAATTTGGGGGTGATGTTGTGGGTGGGGTGAACAACAGGCTTACCTTCCACCGTGGT 350
241 GTTATTTGCTAACAGAGAGCAATTTTCATGGGCTAACAAAGACAGATTCACACAGCAAT 300
351 GTTATTTGCTAACAGAGAGCAATTTTCATGGGCTAACAAAGACAGATTCACACAGCAAT 410
301 TACTTTGGCTCCCTCAACCCAGCCTCAGCTATTTGAGTTGATTCCTTCCAGGAGAGGAA 360
411 TACTTTGGCTCCCTCAACCCAGCCTCAGCTATTTGAGTTGATTCCTTCCAGGAGAGGAA 470
361 ATCTATGCCCATTTCAAGAGTCTGCTTCAATGGTTAATCTCGAGACATTTGTGTTGGG 420
```

Db 471 ATCTATGCCCATTTCAAGAGTCTGCTTCAATGGTTAATCTCGAGACATTTGTGTTGG 530
Qy 421 GGATGGATATACAGCAACATGAACTGGCTGATGGCATGGCCAGGACAA 470
Db 531 GGATGGATATACAGCAACATGAACTGGCTGATGGCATGGCCAGGACAA 580

RESULT 13
Bg652636
LOCUS
DEFINITION
Bg652636 516 bp mRNA linear EST 22-JUL-2004
ID: Gm-cl051-5668 5' similar to TR:Q9SSV4 Q9SSV4
MTO-INOSITOL-1-PHOSPHATE SYNTHASE.; mRNA sequence.

ACCESSION

VERSION

Bg652636.1 GI:13790045

EST.

KEYWORDS

SOURCE

Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE

1 (bases 1 to 516)
Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kueba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
McCam, R., Riter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCam, R., Waterston, R. and Wilson, R.

TITLE
JOURNAL
COMMENT

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
High quality sequence stop: 383.

FEATURES

source

1. 516
Location/Qualifiers

/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Corolla"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-5668"
/tissue_type="floral meristematic mRNA"
/lab_host="DH10B"
/clone_lib="Gm-cl051"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from floral
meristematic mRNA provided by Dr. Halina Knop of Clemson
University. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 26.9%; Score 413; DB 4; Length 516;
Best Local Similarity 99.6%; Pred. No. 3.6e-225;
Matches 513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
812 GTGTGATGGAATAATGTTCTTTCATTTAGGAAGCCCTCAACAACCTTTGTACCAAGGC 871
```


Db 2 GTGTGATGAAAAATGTTCTTTCATTATGAAAGCCCTCAGAACACTTTTGTACAGAGGC 61
 QY 872 TGAATTATCTTGGCATCGGAGAGAACACTTTGATTGTGAGATGACTTCAAGATGCTC 931
 Db 62 TGAATTATCTTGGCATCGGAGAGAACACTTTGATTGTGAGATGACTTCAAGATGCTC 121
 QY 932 AGACCAAAATGAAATCTGTGTGATTCTTCTGTGGGGGCTGGATCAAGCAACAT 991
 Db 122 AGACCAAAATGAAATCTGTGTGATTCTTCTGTGGGGGCTGGATCAAGCAACAT 181
 QY 992 CTATAGTATGATCAACATCTTGGGAAACATGATGATGATCTCTCGCTCCACAAA 1051
 Db 182 CTATAGTATGATCAACATCTTGGGAAACATGATGATGATCTCTCGCTCCACAAA 241
 QY 1052 CCTTCGGCTCCAGAGAAATCTCCAGAGCACTTTGTGACATATGATCAACAGCAATG 1111
 Db 242 CCTTCGGCTCCAGAGAAATCTCCAGAGCACTTTGTGACATATGATCAACAGCAATG 301
 QY 1112 CCATCTCTATAGAGCTGGTGAACATCCGACCATGTGTGTTATTAAGTATGCTT 1171
 Db 302 CCATCTCTATAGAGCTGGTGAACATCCGACCATGTGTGTTATTAAGTATGCTT 361
 QY 1172 AGTAGAGGATGACAGAGAGAGCCATGATGATGATCACTTCAAGATATTCATGAGTGA 1231
 Db 362 AGTAGAGGATGACAGAGAGAGCCATGATGATGATCACTTCAAGATATTCATGAGTGA 421
 QY 1232 AGAACCCATTTGTTTGACACACATGTGAGATTCCTTTAGTGTCTCTATTATCT 1291
 Db 422 AGAACCCATTTGTTTGACACACATGTGAGATTCCTTTAGTGTCTCTATTATCT 481
 QY 1292 TGGACTTGGTCTCTTCTGCTGAGCTGAGCACTAGA 1326
 Db 482 TGGACTTGGTCTCTTCTGCTGAGCTGAGCACTAGA 516

RESULT 14
 BE660322 505 bp mRNA linear EST 06-SEP-2000
 LOCUS BE660322 Glycine max cDNA, mRNA sequence.
 DEFINITION BE660322.1 GI:9986214
 VERSION EST.
 KEYWORDS
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 505)
 AUTHORS Harris, N., Chapman, B. P. and Gijzen, M.
 TITLES Gene expression in developing soybean seed coats
 JOURNAL Unpublished (2000)
 COMMENT Contact: Gijzen M
 Agriculture and Agri-Food Canada
 1391 Sandford Street, London, Ontario, Canada N5V 4T3
 Tel: 519 457 1470
 Fax: 519 457 3997
 Email: gijzenm@agr.gc.ca.
 Location/Qualifiers
 1..505

FEATURES
 source
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Harosoy 63"
 /db_xref="taxon:3847"
 /tissue_type="Seed coats"
 /lab_host="E. coli strain XL0LR"
 /clone_lib="GmaxSC"
 /note="Vector: pBK-CMV; Site 1: EcoRI, Site 2: XhoI. This
 cDNA library was constructed from polyA+ enriched mRNA
 from green seed coats in mid to late developmental
 stage, average fresh weight 250 mg per seed. Traces of
 pod and embryo tissue also present. Complementary DNA was
 synthesized from mRNA using an XhoI-poly(dT)

linker-primer. EcoRI adapters were ligated to the
 blunt-ended cDNA fragments and the products were digested
 with XhoI for directional cloning into lambda ZAP Express
 vector. This lambda library was amplified once using E.
 coli host strain XL1 Blue MRF+. Inserts were then
 subcloned by mass excision using ExAssist helper phage for
 conversion into phagemid vector pBK-CMV in E. coli host
 strain XL0LR."

Query Match 24.6%; Score 377; DB 2; Length 505;
 Best local Similarity 99.8%; Pred. No. 1,9e-204;
 Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTTCATGAGAAATTTTAAGGTAGAGAGTCTTATGTGAAGTACCCAGACTGAGATT 60
 Db 77 ATGTTTCATGAGAAATTTTAAGGTAGAGAGTCTTATGTGAAGTACCCAGACTGAGATT 136
 QY 61 CAGTCGGTGTACAACTACGAAACCAACCCGAACTTTGTACAGAGAAACGAAATGGCACTTAT 120
 Db 137 CAGTCGGTGTACAACTACGAAACCAACCCGAACTTTGTACAGAGAAACGAAATGGCACTTAT 196
 QY 121 CAGTGGATTTGCAAAACCCAAATCCGTCACATCAATTTAAACCAACCAACCACTTTCCA 180
 Db 197 CAGTGGATTTGCAAAACCCAAATCCGTCACATCAATTTAAACCAACCAACCACTTTCCA 256
 QY 181 AAATTGGGGGTGATGCTTGTGGGTTGGGGTGAACCAACGCTCTACACCGGTGCT 240
 Db 257 AAATTGGGGGTGATGCTTGTGGGTTGGGGTGAACCAACGCTCTACACCGGTGCT 316
 QY 241 GTTATTGCTCAAGAGAGAGACATTTCAAGGGCTACAGAGAAAGAAATTCAGAGCAAT 300
 Db 317 GTTATTGCTCAAGAGAGAGACATTTCAAGGGCTACAGAGAAAGAAATTCAGAGCAAT 376
 QY 301 TACTTGGCTCCCTACACCAACCTCAGCTATTCAGATTCGATTCCTTCCAGGAGAGAGAA 360
 Db 377 TACTTGGCTCCCTACACCAACCTCAGCTATTCGATTCGATTCCTTCCAGGAGAGAGAA 436
 QY 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCATGTTAATCCTGACGACATTTGTGTTGGG 420
 Db 437 ATCTATGCCCATTCAGAGAGTCTGCTTCCATGTTAATCCTGACGACATTTGTGTTGGG 496
 QY 421 GGATGGGA 428
 Db 497 GGATGGGA 504

RESULT 15
 BUS49101/c 588 bp mRNA linear EST 16-SEP-2002
 LOCUS BUS49101 Glycine max cDNA clone Gm-r1088-6667 3',
 DEFINITION BUS49101.1 GI:22931962
 ACCESSION BUS49101
 VERSION EST.
 KEYWORDS
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 588)
 AUTHORS Vodkin, L., Shoemaker, R., Keim, P., Retzel, E., Khanna, A., Shealy, R.,
 Clough, S., Thibaud-Nissen, F., Coryell, V., Erpelid, J., Raph, C.,
 Shoop, E., Stromvik, M., Schweitzer, P., Gong, G. and Liu, L.
 TITLES A Functional Genomics Program for Soybean (NSF 9872565) (2002)
 JOURNAL Unpublished (2002)
 COMMENT Other ESTs: BF324641 corresponding to Gm-c1066-1370 (5')
 Contact: Vodkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics
 University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vodk@uiuc.edu
 Insert Length: 588 Std Error: 0.00
 Plate: GM880018B10 row: C column: 10
 Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'
 High quality sequence stop: 588.
 Location/Qualifiers
 1. 588
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone_lib="Gm-r1088-6667"
 /clone_lib="Gm-r1088"
 /note="The library Gm-r1088 is a sequence-driven, reracked set of 9,216 low redundancy clones selected from cDNA libraries from various tissues and stages of development of soybean that consists of 2,706 cDNAs from germinating cotyledons (source library Gm-c1027); 1,355 cDNAs from immature seed coats (libraries Gm-c1019 and Gm-c1023); 917 cDNAs from tissue culture derived somatic embryos (source libraries Gm-c1036 and Gm-c1075); 3,275 cDNAs from germinating seedlings, shoot tips, or leaves exposed to various stresses (source libraries Gm-c1064, Gm-c1065, Gm-c1066, and Gm-c1067; and Gm-c1068); and 963 cDNAs from young leaves exposed to bacterial and fungal pathogens (source libraries Gm-c1072, Gm-c1073, and Gm-c1074). The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1088 and the cDNA clones of the reracked Gm-r1088 library were then sequenced at the 3' end. The unigene selection and 3' sequencing was funded by NSF Plant Genome project #9872565 (http://soybean.genomics.croplsci.uiuc.edu/) as part of creation of a low redundancy soybean cDNA set. The source cDNA libraries were constructed by the laboratories of Lila Vodkin, University of Illinois, Randy Shoemaker, Iowa State University, and Paul Keim, Northern Arizona University as part of the Public EST project. http://129.186.26.94/soybeanest.html. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nfs0y/. Reracking and 3' sequencing were conducted by services of the University of Illinois Beck Center for Comparative and Functional Genomics http://www.biotech.uiuc.edu/beck.htm. Note: The corresponding 5' EST from each clone in the Gm-r1088 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 24.4%; Score 374; DB 5; Length 588;
 Best Local Similarity 100.0%; Pred. No. 1e-202;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1160 AGTATGTGCTTACGTAGGGATAGCAAGAGCCATGATGATACCTTCAAGATAT 1219
 DB 501 AGTATGTGCTTACGTAGGGATAGCAAGAGCCATGATGATACCTTCAAGATAT 442
 QY 1220 TCATGGGTGGAAGAACACCATTTGTTGCACAACACATGTGAGATTCCCTTTAGCTG 1279
 DB 441 TCATGGGTGGAAGAACACCATTTGTTGCACAACACATGTGAGATTCCCTTTAGCTG 382
 QY 1280 CTCCTATTATCTTGGACTTGTCCTCTGAGCTGAGCACTAGATCCAGTTTAAAG 1339
 DB 381 CTCCTATTATCTTGGACTTGTCCTCTGAGCTGAGCACTAGATCCAGTTTAAAG 322
 QY 1340 CTGAAATGAGGGAATTCACATTCACCCAGTTGCTACATTCACGTAATCTGA 1399

DB 321 CTGAAATGAGGGAATTCACATTCACCCAGTTGCTACATTCACGTAATCTGA 262
 QY 1400 CCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGTGATGATTTGCAAGACGCTG 1459
 DB 261 CCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGTGATGATTTGCAAGACGCTG 202
 QY 1460 CAATGCTGGAACATTAATGAGGGCTTGTGTTGATTTGGCCCAAGAAATACATGATTC 1519
 DB 201 CAATGCTGGAACATTAATGAGGGCTTGTGTTGATTTGGCCCAAGAAATACATGATTC 142
 QY 1520 TCGAGTACAGTGA 1533
 DB 141 TCGAGTACAGTGA 128

Search completed: June 8, 2005, 14:33:48
 Job time : 5225.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:50:16 ; Search time 260.698 Seconds
(without alignments)
9621.891 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgctcctcgcagattttaa.....tgattctcgcagcacaagtcga 1533

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/6CTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile01.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	23	1.5	1931	3	US-09-118-442-10
2	23	1.5	1931	3	US-09-677-064-10
3	23	1.5	3546	3	US-09-118-442-14
4	23	1.5	3546	3	US-09-118-442-15
5	23	1.5	3546	3	US-09-677-064-14
6	23	1.5	3546	3	US-09-677-064-15
7	20	1.3	27223	4	US-09-949-016-13036
8	19	1.2	1258	1	US-08-539-798-1
9	19	1.2	1258	1	US-08-329-560-1
10	19	1.2	3153	3	US-09-175-928-9
11	18	1.2	256	4	US-09-313-294A-2055
12	18	1.2	271	4	US-09-513-999C-9342
13	18	1.2	472	3	US-09-641-638-549
14	18	1.2	472	3	US-09-641-638-550
15	18	1.2	472	4	US-10-170-097-549
16	18	1.2	472	4	US-10-170-097-550
17	18	1.2	601	4	US-09-949-016-41180
18	18	1.2	601	4	US-09-949-016-112271
19	18	1.2	601	4	US-09-949-016-141642
20	18	1.2	1001	3	US-09-641-638-278
21	18	1.2	1001	4	US-10-170-097-278
22	18	1.2	1398	4	US-09-328-352-429
23	18	1.2	2477	4	US-09-907-794A-169
24	18	1.2	2477	4	US-09-905-125A-169
25	18	1.2	2477	4	US-09-902-775A-169
26	18	1.2	2477	4	US-09-906-700-169
27	18	1.2	2477	4	US-09-903-603A-169

c 28	18	1.2	2477	4	US-09-904-920A-169	Sequence 169, App
c 29	18	1.2	2477	4	US-09-909-064-169	Sequence 169, App
c 30	18	1.2	2477	4	US-09-905-381A-169	Sequence 169, App
c 31	18	1.2	2477	4	US-09-906-618A-169	Sequence 169, App
c 32	18	1.2	4989	4	US-09-693-011-12	Sequence 12, Appl
c 33	18	1.2	5083	4	US-09-693-011-11	Sequence 11, Appl
c 34	18	1.2	6314	4	US-09-693-011-10	Sequence 10, Appl
c 35	18	1.2	6408	4	US-09-693-011-9	Sequence 9, Appl
c 36	18	1.2	9060	3	US-08-378-313-20	Sequence 20, Appl
c 37	18	1.2	16573	4	US-09-949-016-11764	Sequence 11764, A
c 38	18	1.2	17370	4	US-09-949-016-11731	Sequence 1731, A
c 39	18	1.2	55264	4	US-09-949-016-15014	Sequence 15014, A
c 40	18	1.2	56147	4	US-09-949-016-16352	Sequence 16352, A
c 41	18	1.2	87617	4	US-09-949-016-16551	Sequence 16551, A
c 42	18	1.2	92227	4	US-09-949-016-11929	Sequence 11929, A
c 43	18	1.2	92232	4	US-09-949-016-15421	Sequence 15421, A
c 44	18	1.2	94755	4	US-09-949-016-11839	Sequence 11839, A
c 45	18	1.2	107140	4	US-09-949-016-14834	Sequence 14834, A

ALIGNMENTS

```
RESULT 1
US-09-118-442-10
Sequence 10, Application US/09118442B
Patent No. 6197561
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
APPLICANT: Wang, Xun
APPLICANT: Bowen, Benjamin A.
TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
FILE REFERENCE: 0706
CURRENT APPLICATION NUMBER: US/09/118,442B
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: 60/055,446
EARLIER FILING DATE: 1997-08-11
EARLIER APPLICATION NUMBER: 60/055,526
EARLIER FILING DATE: 1997-08-08
EARLIER APPLICATION NUMBER: 60/053,944
EARLIER FILING DATE: 1997-07-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 1931
TYPE: DNA
ORGANISM: Zea mays
US-09-118-442-10
Query Match 1.5%; Score 23; DB 3; Length 1931;
Best local Similarity 100.0%; Pred. No. 0.22;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1105 AGCAATGCCATCTCTATGAGCC 1127
DB 1203 AGCAATGCCATCTCTATGAGCC 1225
RESULT 2
US-09-677-064-10
Sequence 10, Application US/09677064
Patent No. 6291224
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
FILE REFERENCE: 0706D
CURRENT APPLICATION NUMBER: US/09/677,064
```

```
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-10

Query Match          1.5%; Score 23; DB 3; Length 1931;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1105 AGCAATGCCATCCTCTATGAGCC 1127
      |||
      1203 AGCAATGCCATCCTCTATGAGCC 1225

RESULT 3
US-09-118-442-14
; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-14

Query Match          1.5%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1105 AGCAATGCCATCCTCTATGAGCC 1127
      |||
      2842 AGCAATGCCATCCTCTATGAGCC 2864

RESULT 4
US-09-118-442-15
; Sequence 15, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
```

```
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-15

Query Match          1.5%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1105 AGCAATGCCATCCTCTATGAGCC 1127
      |||
      2842 AGCAATGCCATCCTCTATGAGCC 2864

RESULT 5
US-09-677-064-14
; Sequence 14, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-14

Query Match          1.5%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1105 AGCAATGCCATCCTCTATGAGCC 1127
      |||
      2842 AGCAATGCCATCCTCTATGAGCC 2864

RESULT 6
US-09-677-064-15
; Sequence 15, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
```

```

; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
; US-09-677-064-15

Query Match          1.5%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1105 AGCAATGCCATCTCTATAGCC 1127
Db      2842 AGCAATGCCATCTCTATAGCC 2864

RESULT 7
US-09-949-016-13036
; Sequence 13036, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13036
; LENGTH: 27223
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-13036

Query Match          1.3%; Score 20; DB 4; Length 27223;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1476 AATGAGGCTGTGTGAT 1495
Db      3336 AATGAGGCTGTGTGAT 3355

RESULT 8
US-08-539-798-1
; Sequence 1, Application US/08539798
; Patent No. 5614400
; GENERAL INFORMATION:
; APPLICANT: CAHOON, Edgar B.
; APPLICANT: OHROGEE, John B.
; APPLICANT: OHROGEE, John B.
```

```

; TITLE OF INVENTION: Methods and Compositions Relating to
; TITLE OF INVENTION: Plant 6-Delta Palmitoyl-Acyl Carrier Protein Desaturase
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pioneer Hi-Bred International, Inc.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: US
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,798
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,560
; FILING DATE: 26-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Yates, Michael B.
; REGISTRATION NUMBER: 36,063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1258 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-539-798-1

Query Match          1.2%; Score 19; DB 1; Length 1258;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      327 AGCTATCGAGTTGATCC 345
Db      819 AGCTATCGAGTTGATCC 837

RESULT 9
US-08-329-560-1
; Sequence 1, Application US/08329560
; Patent No. 5654402
; GENERAL INFORMATION:
; APPLICANT: CAHOON, Edgar B.
; APPLICANT: OHROGEE, John B.
; TITLE OF INVENTION: Methods and Compositions Relating to
; TITLE OF INVENTION: Plant 6-Delta Palmitoyl-Acyl Carrier Protein Desaturase
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pioneer Hi-Bred International, Inc.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: US
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/329,560
; FILING DATE: 26-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

NAME: Yates, Michael E.
REGISTRATION NUMBER: 36,063
REFERENCE/DOCKET NUMBER: 0284US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1258 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-329-560-1

Query Match 1.2%; Score 19; DB 1; Length 1258;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 AGCTATTCGAGTTGGATCC 345
|||
Db 819 AGCTATTCGAGTTGGATCC 837

RESULT 10
US-09-175-928-9/c
Sequence 9, Application US/09175928A
Patent No. 6312921
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John W.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Mi, Sha
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6006B.AJ172A
CURRENT APPLICATION NUMBER: US/09/175,928A
CURRENT FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 3153
TYPE: DNA
ORGANISM: Homo sapiens
US-09-175-928-9

Query Match 1.2%; Score 19; DB 3; Length 3153;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1100 TCACACGATGCCATCCT 1118
|||
Db 178 TCACACGATGCCATCCT 160

RESULT 11
US-09-313-294A-2055
Sequence 2055, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2055

LENGTH: 256
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700551970H1
US-09-313-294A-2055

Query Match 1.2%; Score 18; DB 4; Length 256;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 CAACATGAACCTGGCTGA 452
|||
Db 99 CAACATGAACCTGGCTGA 116

RESULT 12
US-09-513-999C-9342
Sequence 9342, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.
APPLICANT: Duclet, A.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59.US2.REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 9342
LENGTH: 271
TYPE: DNA
ORGANISM: Homo sapiens
US-09-513-999C-9342

Query Match 1.2%; Score 18; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 TCCTCACCACCAAGCCTCA 327
|||
Db 144 TCCTCACCACCAAGCCTCA 161

RESULT 13
US-09-641-638-549
Sequence 549, Application US/09641638
Patent No. 6432648
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bouguetelret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLULIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GENSET.051CP1
CURRENT APPLICATION NUMBER: US/09/641,638
CURRENT FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm

```
; SEQ ID NO 549
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 416
; OTHER INFORMATION: 12-63-402 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 396..415
; OTHER INFORMATION: 12-63-402.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 417..436
; OTHER INFORMATION: 12-63-402.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 15..35
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 445..464
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 404..428
; OTHER INFORMATION: 12-63-402 potential probe
; NAME/KEY: misc_feature
; LOCATION: 13..14
; OTHER INFORMATION: n=a, g, c or t
; US-09-641-638-549
```

```
Query Match 1.2%; Score 18; DB 3; Length 472;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 767 GAAATGAGCTGAGATT 784
Db 199 GAAATGAGCTGAGATT 216
```

```
RESULT 14
US-09-641-638-550
; Sequence 550, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 550
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88
; OTHER INFORMATION: 12-63-74 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 68..87
; OTHER INFORMATION: 12-63-74.misl, potential
; NAME/KEY: misc_binding
```

```
; LOCATION: 89..108
; OTHER INFORMATION: 12-63-74.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 15..35
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 445..464
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 76..100
; OTHER INFORMATION: 12-63-74 potential probe
; NAME/KEY: misc_feature
; LOCATION: 13..14
; OTHER INFORMATION: n=a, g, c or t
; US-09-641-638-550
```

```
Query Match 1.2%; Score 18; DB 3; Length 472;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 767 GAAATGAGCTGAGATT 784
Db 199 GAAATGAGCTGAGATT 216
```

```
RESULT 15
US-10-170-097-549
; Sequence 549, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T14XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 549
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 416
; OTHER INFORMATION: 12-63-402 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 396..415
; OTHER INFORMATION: 12-63-402.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 417..436
; OTHER INFORMATION: 12-63-402.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 15..35
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
```



```

: LOCATION: 445..464
: OTHER INFORMATION: downstream amplification primer, complement
:
: FEATURE:
: NAME/KEY: misc_binding
: LOCATION: 404..428
: OTHER INFORMATION: 12-63-402 potential probe
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 13..14
: OTHER INFORMATION: n=a, g, c or t
:
US-10-170-097-549

```

Query Match	1.2%;	Score 18;	DB 4;	Length 472;
Best Local Similarity	100.0%;	Pred. No. 79;		
Matches 18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	767	GAATGAGCGTGAGATT	784
Db	199	GAATGAGCGTGAGATT	216

Search completed: June 8, 2005, 14:43:08
Job time : 260.698 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 11:26:41 ; Search time 950.618 Seconds
(without alignments)
9996.820 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 aagttcaccgagaatttaa.....tgattctcgagtaacgta 1533

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 6046767 seqs, 3099530249 residues

Word size : 0

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17:	/cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18:	/cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19:	/cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20:	/cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21:	/cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22:	/cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq.*
23:	/cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1533	100.0	1533	14	US-10-025-003-11
2	1533	100.0	1533	14	US-10-718-952-11
3	1482	96.7	1533	14	US-10-025-003-13
4	1482	96.7	1533	14	US-10-025-003-15
5	1482	96.7	1533	14	US-10-718-952-13
6	1482	96.7	1533	19	US-10-718-952-15
7	986	64.3	2582	18	US-10-424-599-12021

8	310	20.2	634	18	US-10-424-599-22663	Sequence 22663, A
9	185	12.1	844	18	US-10-424-599-70162	Sequence 70162, A
10	146	9.5	661	18	US-10-424-599-116597	Sequence 116597, A
11	146	9.5	1533	14	US-10-025-003-5	Sequence 5, Appl1
12	146	9.5	1533	14	US-10-025-003-9	Sequence 9, Appl1
13	146	9.5	1533	14	US-10-718-952-5	Sequence 5, Appl1
14	146	9.5	1533	19	US-10-718-952-9	Sequence 9, Appl1
15	146	9.5	1760	14	US-10-025-003-1	Sequence 1, Appl1
16	146	9.5	1760	14	US-10-718-952-1	Sequence 1, Appl1
17	146	9.5	1589	18	US-10-424-599-70167	Sequence 70167, A
18	121	7.9	594	18	US-10-424-599-70166	Sequence 70166, A
19	44	2.9	2018	18	US-10-424-599-12022	Sequence 12022, A
20	35	2.3	989	18	US-10-424-599-12019	Sequence 12019, A
21	35	2.3	2582	18	US-10-424-599-12021	Sequence 12021, A
22	30	2.0	535	18	US-10-424-599-78882	Sequence 78882, A
23	28	1.8	355	18	US-10-424-599-94906	Sequence 94906, A
24	27	1.8	252	18	US-10-424-599-12017	Sequence 12017, A
25	26	1.7	377	11	US-09-733-6277A-2381	Sequence 2381, Ap
26	26	1.7	532	19	US-10-021-323-16419	Sequence 16419, A
27	26	1.7	533	19	US-10-021-323-16951	Sequence 16951, A
28	26	1.7	543	19	US-10-021-323-16651	Sequence 16651, A
29	26	1.7	566	19	US-10-021-323-12655	Sequence 12655, A
30	26	1.7	618	19	US-10-021-323-17771	Sequence 17771, A
31	26	1.7	1673	20	US-10-425-115-47180	Sequence 47180, A
32	26	1.7	1941	20	US-10-739-930-2091	Sequence 2091, Ap
33	24	1.6	451	18	US-10-424-599-137655	Sequence 137655, A
34	24	1.6	1924	14	US-10-767-701-14714	Sequence 14714, A
35	23	1.5	35	19	US-10-025-003-3	Sequence 3, Appl1
36	23	1.5	35	19	US-10-718-952-3	Sequence 3, Appl1
37	23	1.5	260	18	US-10-424-599-36573	Sequence 36573, A
38	23	1.5	290	19	US-10-437-963-16182	Sequence 16182, A
39	23	1.5	319	18	US-10-424-599-106050	Sequence 106050, A
40	23	1.5	345	11	US-09-733-6277A-3508	Sequence 3508, Ap
41	23	1.5	452	11	US-09-733-6277A-3572	Sequence 3572, A
42	23	1.5	557	19	US-10-021-323-12840	Sequence 12840, A
43	23	1.5	601	19	US-10-021-323-15202	Sequence 15202, A
44	23	1.5	605	19	US-10-021-323-15202	Sequence 15202, A
45	23	1.5	1898	18	US-10-425-114-7081	Sequence 7081, Ap

ALIGNMENTS

RESULT 1
US-10-025-003-11
; Sequence 11, Application US/10025003
; Publication No. US20030074685A1
GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-11
Query Match 100.0%; Score 1533; DB 14; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGTCATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAGTACACCGAGACTGAGATT 60
DB 1 AATGTCATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAGTACACCGAGACTGAGATT 60
OY 61 CAGTCGGTGTACACCTACGAAACCAACGAACTTTGTTCAGAGAAACAGAAATGGCACCTT 120
DB 61 CAGTCGGTGTACACCTACGAAACCAACGAACTTTGTTCAGAGAAACAGAAATGGCACCTT 120
OY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACCTACCAATTTAAAAACCAACCCGATTTCCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACCTACCAATTTAAAAACCAACCCGATTTCCA 180
OY 181 AAATTTGGGGGTATGCTTGTGGGTGGGGTGAACCAACGGCTCTACCTCAACGGTGT 240
DB 181 AAATTTGGGGGTATGCTTGTGGGTGGGGTGAACCAACGGCTCTACCTCAACGGTGT 240
OY 241 GTTATTTGCTAACAGAGAGACATTTATGTGGCTACAAAGACAAAGATTCAACAGCCAT 300
DB 241 GTTATTTGCTAACAGAGAGACATTTATGTGGCTACAAAGACAAAGATTCAACAGCCAT 300
OY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTCAAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTCAAGGAGAGAA 360
OY 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATGTGTTAATCTGACGACATTTGTTGG 420
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATGTGTTAATCTGACGACATTTGTTGG 420
OY 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGCCATGGCCAGGGCAAAAGTGTTCAC 480
DB 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGCCATGGCCAGGGCAAAAGTGTTCAC 480
OY 481 ATCGATTTTGCAGAGAGATTTGAGGCTTACATGGAATCCATGGTTCATCCCGGAAATC 540
DB 481 ATCGATTTTGCAGAGAGATTTGAGGCTTACATGGAATCCATGGTTCATCCCGGAAATC 540
OY 541 TAGCAACCCGAGATTTCTTGTCTGCAACCAAGAGAGCGTCCCAACACGTGATTAAGG 600
DB 541 TAGCAACCCGAGATTTCTTGTCTGCAACCAAGAGAGCGTCCCAACACGTGATTAAGG 600
OY 601 ACAAAAGCAAGACAGTTCAGCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACC 660
DB 601 ACAAAAGCAAGACAGTTCAGCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACC 660
OY 661 AAAGTGAACAAGGTGTTGCTGCTGAGACTGCCAACAAGAGAGGTATGCAATTTGTT 720
DB 661 AAAGTGAACAAGGTGTTGCTGCTGAGACTGCCAACAAGAGAGGTATGCAATTTGTT 720
OY 721 GTAGGCTTATGACACCATGAGAACTCTTGTGCTGTGAGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTATGACACCATGAGAACTCTTGTGCTGTGAGACAGAAATGAGGCTGAG 780
OY 781 ATTTCTCTTCCAACCTTTGATGCAATGCTGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
DB 781 ATTTCTCTTCCAACCTTTGATGCAATGCTGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
OY 841 GGAAGCCCTCAGAAACCTTTTGAACAGAGGCTGATGATCTTGCCATGCGAGAGAACT 900
DB 841 GGAAGCCCTCAGAAACCTTTTGAACAGAGGCTGATGATCTTGCCATGCGAGAGAACT 900
OY 901 TTGATTTGTGAGATGACTTCAAGAGTGTGAGACCAAAATGAATCTGTGTTGTTGAT 960
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGAGACCAAAATGAATCTGTGTTGTTGAT 960
OY 961 TTTCTTTGGGGGCTGGTATCAAGCAACATCTAATGTTAGTTAACAACATCTGGGAAAC 1020
DB 961 TTTCTTTGGGGGCTGGTATCAAGCAACATCTAATGTTAGTTAACAACATCTGGGAAAC 1020
OY 1021 AATGATGATGATGATGATCTCGGCTCAAAACCTTCCGCTCAAGAGAAATCTTCAAGAAC 1080
DB 1021 AATGATGATGATGATGATCTCGGCTCAAAACCTTCCGCTCAAGAGAAATCTTCAAGAAC 1080
OY 1081 AAGCTTGTGAGATGATGATGATCAACAGCAATGCAATCCTATGAGCGCTGGTGAACATCCC 1140
```

```
DB 1081 AACGTTGTGACGATATGATCAACAGCAATGCATCTCTATGAGCCCTGGTGAACATCCC 1140
OY 1141 GACCATGTTGTTGTTAATTAAGTAATGTCCTTAACGTAAGGGGATAGCAAGAGCCATGAT 1200
DB 1141 GACCATGTTGTTGTTAATTAAGTAATGTCCTTAACGTAAGGGGATAGCAAGAGCCATGAT 1200
OY 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAAGAACCAACATTTGTTTTCACAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAAGAACCAACATTTGTTTTCACAACATGT 1260
OY 1261 GAGGATTCCTTTTACGCTGCTCTATTAATTTTGAACCTTGTCTTCTTGTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTACGCTGCTCTATTAATTTTGAACCTTGTCTTCTTGTGAGCTGAGC 1320
OY 1321 ACTAATATCCAGTTTAAAGCTGAAATGAGGAAATTCACCTCATTCACCCAGTTGCT 1380
DB 1321 ACTAATATCCAGTTTAAAGCTGAAATGAGGAAATTCACCTCATTCACCCAGTTGCT 1380
OY 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTGTTCCACCGGGTACACAGTGTGAT 1440
DB 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTGTTCCACCGGGTACACAGTGTGAT 1440
OY 1441 GCATTGTCAAAAGCAGCTGCAATGCTGAAACATTAAGAGGCTTGTGTTGATGGCC 1500
DB 1441 GCATTGTCAAAAGCAGCTGCAATGCTGAAACATTAAGAGGCTTGTGTTGATGGCC 1500
OY 1501 CCAGAGAAATTAACATGATCTCGAGTACAAAGTGA 1533
DB 1501 CCAGAGAAATTAACATGATCTCGAGTACAAAGTGA 1533
```

RESULT 2

US-10-718-952-11

Sequence 11, Application US/10718952
Publication No. US20040128713A1

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

FILE REFERENCE: BB-1077-C

CURRENT APPLICATION NUMBER: US/10/718,952

PRIORITY FILING DATE: 2003-11-21

PRIORITY FILING DATE: APRIL 8, 1997

PRIORITY FILING DATE: APRIL 7, 1998

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97

SEQ ID NO 11

LENGTH: 1533

TYPE: DNA

ORGANISM: Glycine max

US-10-718-952-11

Query Match 100.0%; Score 1533; DB 19; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 AATGTCATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAGTACACCGAGACTGAGATT 60
DB 1 AATGTCATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAGTACACCGAGACTGAGATT 60
OY 61 CAGTCGGTGTACACCTACGAAACCAACGAACTTTGTTCAGAGAAACAGAAATGGCACCTT 120
DB 61 CAGTCGGTGTACACCTACGAAACCAACGAACTTTGTTCAGAGAAACAGAAATGGCACCTT 120
OY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACCTACCAATTTAAAAACCAACCCGATTTCCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACCTACCAATTTAAAAACCAACCCGATTTCCA 180
```

```
QY 181 AATTTGGGGTGTATGCTTGTGGGTGGGGAACAACGGCTTACCCCTCACCGGTGT 240
Db 181 AATTTGGGGTGTATGCTTGTGGGTGGGGAACAACGGCTTACCCCTCACCGGTGT 240
QY 241 GTTATTGCTTAACAGAGACATTTTCAATGGGCTACAAAGACAAGATTCAACAGCAAT 300
Db 241 GTTATTGCTTAACAGAGACATTTTCAATGGGCTACAAAGACAAGATTCAACAGCAAT 300
QY 301 TACTTTGGCTCCCTCACCCCAAGCCTCAGTATTGAGTTGGAATCTTTCAGGAGAGAA 360
Db 301 TACTTTGGCTCCCTCACCCCAAGCCTCAGTATTGAGTTGGAATCTTTCAGGAGAGAA 360
QY 361 ATCTATGCCCCCATTTCAAGAGTCTGCTTCCAAATGGTTAACTCTCAACCAATTTGTTGG 420
Db 361 ATCTATGCCCCCATTTCAAGAGTCTGCTTCCAAATGGTTAACTCTCAACCAATTTGTTGG 420
QY 421 GGATGGGATATCAGCAACATGAACTTGCTGATGCCATGGCCAGGGCAAGGTGTTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACTTGCTGATGCCATGGCCAGGGCAAGGTGTTGAC 480
QY 481 ATCGATTTCGAGAAGCAGTTGAGGCTTACATGGAATCCATGGTTCCACTCCCGGAATC 540
Db 481 ATCGATTTCGAGAAGCAGTTGAGGCTTACATGGAATCCATGGTTCCACTCCCGGAATC 540
QY 541 TAGACACCCGATTTTCATTTGCTGCCAACAAAGAGAGGCTGCCAAACGATTTAAGGCG 600
Db 541 TAGACACCCGATTTTCATTTGCTGCCAACAAAGAGAGGCTGCCAAACGATTTAAGGCG 600
QY 601 ACAAGCAAGAGCAAGTTCAGCAAAATCATCAAGAATCAAGGCGTTTAAAGAGACACC 660
Db 601 ACAAGCAAGAGCAAGTTCAGCAAAATCATCAAGAATCAAGGCGTTTAAAGAGACACC 660
QY 661 AAAAGTGAACAAGGTGTTGCTCTGTGAGCTGCCAACACAGAGAGTATAGCAATTTGGTT 720
Db 661 AAAAGTGAACAAGGTGTTGCTCTGTGAGCTGCCAACACAGAGAGTATAGCAATTTGGTT 720
QY 721 GTTAGGCTTATATACACATGAGGAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
Db 721 GTTAGGCTTATATACACATGAGGAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGTATGCTGCTGATGCTGTGATGGAATAATGTTCTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCTGCTGATGCTGTGATGGAATAATGTTCTTCAATTAAT 840
QY 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATTTGATCTTTCGCAACGAGAACAT 900
Db 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATTTGATCTTTCGCAACGAGAACAT 900
QY 901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
Db 901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
QY 961 TTTCTTGGGGGCTGGTATCAAGCAACATATATAGTTAACAACAATCGGGAAAC 1020
Db 961 TTTCTTGGGGGCTGGTATCAAGCAACATATATAGTTAACAACAATCGGGAAAC 1020
QY 1021 AATGATGTATGAATCTCTCGGCTCCAAACCTTCGCTCAAGAGAAATCTTCAAGAGC 1080
Db 1021 AATGATGTATGAATCTCTCGGCTCCAAACCTTCGCTCAAGAGAAATCTTCAAGAGC 1080
QY 1081 AAGCTTTGTCAGATATGTCACACAGCAATGCAATCTCTATATAGCTGTGTAACATCCC 1140
Db 1081 AAGCTTTGTCAGATATGTCACACAGCAATGCAATCTCTATATAGCTGTGTAACATCCC 1140
QY 1141 GACCATGTTGTTTATTAAGTATGCTTACGTAGGGGATAGCAAGAGCATGGAAT 1200
Db 1141 GACCATGTTGTTTATTAAGTATGCTTACGTAGGGGATAGCAAGAGCATGGAAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAACACATTTGTTGCAACACATGT 1260
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAACACATTTGTTGCAACACATGT 1260
```

```
QY 1261 GAGGATTCCTTTTACGCTCCTATTAATCTTGAATGCTGCTCTGTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTTACGCTCCTATTAATCTTGAATGCTGCTCTGTGAGCTGAGC 1320
QY 1321 ACTAGAAATCAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACCGATTGCT 1380
Db 1321 ACTAGAAATCAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACCGATTGCT 1380
QY 1381 ACCATTCTAGCTATCTGACCAAGGCTCCTCTGCTTCCACCGGGTACACCATGTGGAAT 1440
Db 1381 ACCATTCTAGCTATCTGACCAAGGCTCCTCTGCTTCCACCGGGTACACCATGTGGAAT 1440
QY 1441 GCATTCTCAAGAGGCTGCAATGCTGAAAACATTAAGAGGCTGTGTGGAATGGCC 1500
Db 1441 GCATTCTCAAGAGGCTGCAATGCTGAAAACATTAAGAGGCTGTGTGGAATGGCC 1500
QY 1501 CCAGAGATTAACATGATTTCTCGATCAAGTGA 1533
Db 1501 CCAGAGATTAACATGATTTCTCGATCAAGTGA 1533
```

```
RESULT 3
US-10-025-003-13
; Sequence 13, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-025-003-13
```

```
Query Match 96.7%; Score 1482; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ANTTTCATGAGAATTTTAAAGTAGAGAGTCTTAATGTGAATGACACCGAGCTGAGATT 60
Db 1 ANTTTCATGAGAATTTTAAAGTAGAGAGTCTTAATGTGAATGACACCGAGCTGAGATT 60
QY 61 CAGTCCGTGTACATACGAAACCAACCGAATTTGTTCAAGAGACAGAAATGCGACATTAT 120
Db 61 CAGTCCGTGTACATACGAAACCAACCGAATTTGTTCAAGAGACAGAAATGCGACATTAT 120
QY 121 CAGTGAATGTCAAAACCCAAATCCGTCACATGCAATTTAAACCAACCAACCGATGTCA 180
Db 121 CAGTGAATGTCAAAACCCAAATCCGTCACATGCAATTTAAACCAACCAACCGATGTCA 180
QY 181 AATTTGGGGTATGCTTGTGGGTGGGGAACAACGAGCTTACCTTCAACCGGTGCT 240
Db 181 AATTTGGGGTATGCTTGTGGGTGGGGAACAACGAGCTTACCTTCAACCGGTGCT 240
QY 241 GTTATTGCTTAACAGAGACATTTTCAATGGGCTACAAAGACAAGATTCAACAGCAAT 300
Db 241 GTTATTGCTTAACAGAGAGGCAATTTTCAATGGGCTACAAAGACAAGATTCAACAGCAAT 300
QY 301 TACTTTGGCTCCCTCACCCCAAGCCTCAGTATTGAGTTGGAATCTTTCAGGAGAGAA 360
Db 301 TACTTTGGCTCCCTCACCCCAAGCCTCAGTATTGAGTTGGAATCTTTCAGGAGAGAA 360
```

Db 301 TACTTTGGCTCCCTCAACCCAGCCTCAGCTATTGAGTTGATCCTTCCAGGGAGAGAA 360
Qy ATCTATGCCCATTTGAAGAGTGTGCTTCCAAATGTTAATCCAGAGCAATGTGTTGGG 420
Db 361 ATCTATGCCCATTTGAAGAGTGTGCTTCCAAATGTTAATCCAGAGCAATGTGTTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGCTATGCTGAGGCAAGGAGTTGTAAC 480
Db 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGCTATGCTGAGGCAAGGAGTTGTAAC 480
Qy 481 ATCGATTTGCAAGACAGTTGAGGCTTATCATGAAATCATGTTTCCACTCCCCGGAATC 540
Db 481 ATCGATTTGCAAGACAGTTGAGGCTTATCATGAAATCATGTTTCCACTCCCCGGAATC 540
Qy 541 TACGACCCGGGATTTATGCTGCTGCAACCAAGAGAGCGGCAACAAAGCTGATTAAGAGGC 600
Db 541 TACGACCCGGGATTTATGCTGCTGCAACCAAGAGAGCGGCAACAAAGCTGATTAAGAGGC 600
Qy 601 ACAAGCAAGACAGATTGACGAAATCATCAAAAGACATCAAGCGCTTTAAGAGACCAAC 660
Db 601 ACAAGCAAGACAGATTGACGAAATCATCAAAAGACATCAAGCGCTTTAAGAGACCAAC 660
Qy 661 AAAGTGAACAGGTGTTGCTCTGTGAGTCTGCCAACACAGAGATATGCAATTTGGTT 720
Db 661 AAAGTGAACAGGTGTTGCTCTGTGAGTCTGCCAACACAGAGATATGCAATTTGGTT 720
Qy 721 GTAGGCTTTAATGACCAATGAGAAATCTCTGGGCTGTGAGACAGAAATAGGCTGAG 780
Db 721 GTAGGCTTTAATGACCAATGAGAAATCTCTGGGCTGTGAGACAGAAATAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Qy 841 GGAAGCCTCAGAAACCTTTGTACAGAGGCTGATTTGATCTTCCATCGCAGAACT 900
Db 841 GGAAGCCTCAGAAACCTTTGTACAGAGGCTGATTTGATCTTCCATCGCAGAACT 900
Qy 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGGTGAAT 960
Db 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGGTGAAT 960
Qy 961 TTTCTTGGGGGCTGTGATCAAGCCACATCTATAGTTAGCAACCATCTGGGAAAC 1020
Db 961 TTTCTTGGGGGCTGTGATCAAGCCACATCTATAGTTAGTTAACAACCATCTGGGAAAC 1020
Qy 1021 AATGATGATGATATCTCTGGGCTCAACAACCTTCCGCTCAAGGAAATCTTCAAGAGC 1080
Db 1021 AATGATGATGATATCTCTGGGCTCAACAACCTTCCGCTCAAGGAAATCTTCAAGAGC 1080
Qy 1081 AAGCTTTGAGATATGATCAAGACAGATGCGATCTCTATGAGCGCTGTGAAACATGCC 1140
Db 1081 AAGCTTTGAGATATGATCAAGACAGATGCGATCTCTATGAGCGCTGTGAAACATGCC 1140
Qy 1141 GACCAATGTTGTTAATTAAGTATGCTTACGTAAGGGAATAGCAAGAGCCATGAT 1200
Db 1141 GACCAATGTTGTTAATTAAGTATGCTTACGTAAGGGAATAGCAAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACACATTTGTTTGAACAACATGT 1260
Db 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACACATTTGTTTGAACAACATGT 1260
Qy 1261 GAGGATTCCTTTTAACTGCTCTATTAATCTTGAGCTTGTCTTCTTCTGAGCTGAC 1320
Db 1261 GAGGATTCCTTTTAACTGCTCTATTAATCTTGAGCTTGTCTTCTTCTGAGCTGAC 1320
Qy 1321 ACTAGATATCAGTTTAACTGAAATGAGGGAATTTCCACTATTTCCACCAAGTTGCT 1380
Db 1321 ACTAGATATCAGTTTAACTGAAATGAGGGAATTTCCACTATTTCCACCAAGTTGCT 1380
Qy 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGATACACCACTGCTGAAT 1440
Db 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGATACCACTGCTGAAT 1440

Qy 1441 GCATTGTCAAGACAGCGTGCATATGCTGGAATAATAGAGGCTTGTGATTTGGCC 1500
Db 1441 GCATTGTCAAGACAGCGTGCATATGCTGGAATAATAGAGGCTTGTGATTTGGCC 1500
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACAGTGA 1533
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACAGTGA 1533

RESULT 4
US-10-025-003-15
; Sequence 15, Application US/10025003
; Publication No. US2003074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BR-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-15

Query Match 96.7%; Score 1482; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTTTCATCGAGATTTTAAAGTGAAGTCTCTAATGTGAATACACGAGCTGAGATT 60
Db 1 ATGTTTCATCGAGATTTTAAAGTGAAGTCTCTAATGTGAATACACGAGCTGAGATT 60
Qy 61 CAGTCCTGTCAACATCAAGAAACACCGAATCTTGTTCAGAGAAATGSCACTAT 120
Db 61 CAGTCCTGTCAACATCAAGAAACACCGAATCTTGTTCAGAGAAATGSCACTAT 120
Qy 121 CAGTGAATTTCAAAACCAATCCGTCAACTACAAATTTAAACCAACACCATGTTCCA 180
Db 121 CAGTGAATTTCAAAACCAATCCGTCAACTACAAATTTAAACCAACACCATGTTCCA 180
Qy 181 AAATTGGGGGTATGCTTGTGGTGGGTGGAACAACGGCTCAACCTCAACCGGTGT 240
Db 181 AAATTGGGGGTATGCTTGTGGTGGGTGGAACAACGGCTCAACCTCAACCGGTGT 240
Qy 241 GTTATTGCTAACAGAGAGACATTTATGAGGCTCAAGAGAACAAATTTCAACCAAT 300
Db 241 GTTATTGCTAACAGAGAGAGATTTATGAGGCTCAAGAGAACAAATTTCAACCAAT 300
Qy 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTGAGTTGATGCTTCCAGGAGAGAA 360
Db 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTGAGTTGATGCTTCCAGGAGAGAA 360
Qy 361 ATCTATGCCCATTTGAAGAGTGTGCTTCCAAATGTTAATCCAGAGCAATGTGTTGGG 420
Db 361 ATCTATGCCCATTTGAAGAGTGTGCTTCCAAATGTTAATCCAGAGCAATGTGTTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGCTATGCTGAGGCAAGGAGTTGTAAC 480
Db 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGCTATGCTGAGGCAAGGAGTTGTAAC 480
Qy 481 ATCGATTTGCAAGACAGTTGAGGCTTATCATGAAATCATGTTTCCACTCCCCGGAATC 540

Db 481 ATGATTTGAGAGAGAGTGGAGCCCTTACATGGATTCATGGTTCACCTCCCGGAATC 540
 Qy 541 TAGAGCCCGGATTTCTTGTGCGCAACCAAGAGAGGCTGCCAACACGTGATTAAGGCG 600
 Db 541 TAGAGCCCGGATTTCTTGTGCGCAACCAAGAGAGGCTGCCAACACGTGATTAAGGCG 600
 Qy 601 ACAAGCAAGAGAGAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAGGAAGCACC 660
 Db 601 ACAAGCAAGAGAGAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAGGAAGCACC 660
 Qy 661 AAAGTGAAGAGTGTGTCTGTGAGCTGCAACAGAGAGATTAAGCAATTTGTT 720
 Db 661 AAAGTGAAGAGTGTGTCTGTGAGCTGCAACAGAGAGATTAAGCAATTTGTT 720
 Qy 721 GTAGGCTTAATGACACATGAGAACTCTTGGCTGTGTGACAGAAATGAGCTGAG 780
 Db 721 GTAGGCTTAATGACACATGAGAACTCTTGGCTGTGTGACAGAAATGAGCTGAG 780
 Qy 781 ATTTCTCTTCCACCTTGATGCAATGCTGTGTGATGGAATGTTCTTTCATTAAT 840
 Db 781 ATTTCTCTTCCACCTTGATGCAATGCTGTGTGATGGAATGTTCTTTCATTAAT 840
 Qy 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATGATCTTGCATCGGAGAACACT 900
 Db 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATGATCTTGCATCGGAGAACACT 900
 Qy 901 TTGATTTGAGAGATGATTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
 Db 901 TTGATTTGAGAGATGATTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
 Qy 961 TTTCTTGTGGGGCTGTGATCAAGCCAACTATAGTTAGTTACAAACATCTGGGAAC 1020
 Db 961 TTTCTTGTGGGGCTGTGATCAAGCCAACTATAGTTAGTTAGTTACAAACATCTGGGAAC 1020
 Qy 1021 AATGATGATGATGATCTCGGCTCCAAACCTTCCGCTCCAAAGAAATCTTCAAGAG 1080
 Db 1021 AATGATGATGATGATCTCGGCTCCAAACCTTCCGCTCCAAAGAAATCTTCAAGAG 1080
 Qy 1081 AAGCTTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Db 1081 AAGCTTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Qy 1141 GAGCAGTGTGTTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCT 1200
 Db 1141 GAGCAGTGTGTTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCT 1200
 Qy 1201 GAGTACACTTCAGAGATTTCAATGAGTGGTGAAGAACACATTTGTTGCAACACATGT 1260
 Db 1201 GAGTACACTTCAGAGATTTCAATGAGTGGTGAAGAACACATTTGTTGCAACACATGT 1260
 Qy 1261 GAGGATTCCTTTTAACTGCTCTTATTTCTTGAATGCTTCTTCTTCTGAGCTGAGC 1320
 Db 1261 GAGGATTCCTTTTAACTGCTCTTATTTCTTGAATGCTTCTTCTTCTGAGCTGAGC 1320
 Qy 1321 ACTAAGATTCAGTTTAAAGTATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
 Db 1321 ACTAAGATTCAGTTTAAAGTATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
 Qy 1381 ACCATTCAGCTATGACCAAGGCTCTGCTGATTTCTTGAATGCTTCTTCTGAGCTGAGC 1440
 Db 1381 ACCATTCAGCTATGACCAAGGCTCTGCTGATTTCTTGAATGCTTCTTCTGAGCTGAGC 1440
 Qy 1441 GAGTGTCAAGAGAGGCTGCTGCTGATTTCTTGAATGCTTCTTCTGAGCTGAGC 1500
 Db 1441 GAGTGTCAAGAGAGGCTGCTGCTGATTTCTTGAATGCTTCTTCTGAGCTGAGC 1500
 Qy 1501 CCAAGAGATTAAGATTTCTGAGTCAAGTGA 1533
 Db 1501 CCAAGAGATTAAGATTTCTGAGTCAAGTGA 1533

US-10-718-952-13
 ; Sequence 13, Application US/10718952
 ; Publication No. US20040128713A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hitz, William
 ; APPLICANT: Sebastian, Scott
 ; APPLICANT: Grace, John
 ; APPLICANT: Streitz, Leon
 ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
 ; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
 ; FILE REFERENCE: BB-1077-C
 ; CURRENT APPLICATION NUMBER: US/10/718,952
 ; PRIOR FILING DATE: 2003-11-21
 ; PRIOR APPLICATION NUMBER: 08/835,751
 ; PRIOR FILING DATE: APRIL 8, 1997
 ; PRIOR APPLICATION NUMBER: PCT/US98/06822
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 13
 ; LENGTH: 1533
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-10-718-952-13
 Query Match 96.7%; Score 1482; DB 19; Length 1533;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 ATGTCATGAGAAATTTTAAAGTAGAGTCTTAATGTAAGTACACCGAGCTGAGATT 60
 Db 1 ATGTCATGAGAAATTTTAAAGTAGAGTCTTAATGTAAGTACACCGAGCTGAGATT 60
 Qy 61 CAGTCGTGTACACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGCACTAT 120
 Db 61 CAGTCGTGTACACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGCACTAT 120
 Qy 121 CAGTGTGATGTAACCAACCAACCACTGCACTACCAATTTTAAACCAACCACTGTTCA 180
 Db 121 CAGTGTGATGTAACCAACCAACCACTGCACTACCAATTTTAAACCAACCACTGTTCA 180
 Qy 181 AATTTGGGGGATGATGCTTGTGGGTGGAACCAACGGCTTACCTCAACGGTGT 240
 Db 181 AATTTGGGGGATGATGCTTGTGGGTGGAACCAACGGCTTACCTCAACGGTGT 240
 Qy 241 GTTATGCTACAGAGAGACATTTCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 300
 Db 241 GTTATGCTACAGAGAGACATTTCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 300
 Qy 301 TACTTGGCTCCCTACACCAAGCTTCAAGTATTCAGATTTGATCTTCCAGGAGAGAA 360
 Db 301 TACTTGGCTCCCTACACCAAGCTTCAAGTATTCAGATTTGATCTTCCAGGAGAGAA 360
 Qy 361 ATCTATGCCCATTTGAAGGCTGCTTCAATGTTAATCTGAGAGCAATGTTGTTGG 420
 Db 361 ATCTATGCCCATTTGAAGGCTGCTTCAATGTTAATCTGAGAGCAATGTTGTTGG 420
 Qy 421 GAGTGGATATAGAGACATGAACTGCTGATGCAATGAGTGAAGTGAAGTGAAGTGAAG 480
 Db 421 GAGTGGATATAGAGACATGAACTGCTGATGCAATGAGTGAAGTGAAGTGAAGTGAAG 480
 Qy 481 ATGATTTGACAGAGAGTGAAGCTTCAATGAAATCCATGTTTCACTCCCGGAATC 540
 Db 481 ATGATTTGACAGAGAGTGAAGCTTCAATGAAATCCATGTTTCACTCCCGGAATC 540
 Qy 541 TAGAGCCCGGATTTATGCTGCAACAGAGAGAGGCTGCAACAGAGTGAAGGCG 600
 Db 541 TAGAGCCCGGATTTATGCTGCAACAGAGAGAGGCTGCAACAGAGTGAAGGCG 600
 Qy 601 ACAAGCAAGAGAGAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAGGAAGCACC 660
 Db 601 ACAAGCAAGAGAGAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAGGAAGCACC 660

QY 661 AAAGTGACAGAGTGCTGCTGCTGAGACTGCCAACACAGAGGATGACAAATTTGGTT 720
DB 661 AAAGTGACAGAGTGCTGCTGCTGAGACTGCCAACACAGAGGATGACAAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCAATGAGAAATCTCTGGCTGCTGAGACAGAAATGAGCTGAG 780
DB 721 GTAGGCTTTAATGACACCAATGAGAAATCTCTGGCTGCTGAGACAGAAATGAGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGATGCCATTTGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTGATGCCATTTGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
QY 841 GGAAGCCCTCAGACACTTTTGTACAGAGGCTGATGATCTTGCATGCGAGAACT 900
DB 841 GGAAGCCCTCAGACACTTTTGTACAGAGGCTGATGATCTTGCATGCGAGAACT 900
QY 901 TTGATTTGGTGAGACCTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGAT 960
DB 901 TTGATTTGGTGAGACCTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGAT 960
QY 961 TTTCTTTGGGGCTGATCTCAAGCCAACTATATAGTTAGTACAAACCTGCGGAAAC 1020
DB 961 TTTCTTTGGGGCTGATCTCAAGCCAACTATATAGTTAGTACAAACCTGCGGAAAC 1020
QY 1021 AATGATGATGATATCTCTGAGCTCCAAACCTTCCGCTCCAGAGAAATCTCAAGAGC 1080
DB 1021 AATGATGATGATATCTCTGAGCTCCAAACCTTCCGCTCCAGAGAAATCTCAAGAGC 1080
QY 1081 AACGTTGTTGAGATATGCTCAACAGCAATGCAATCTCTTATAGAGCTGTGTAACATCCC 1140
DB 1081 AACGTTGTTGAGATATGCTCAACAGCAATGCAATCTCTTATAGAGCTGTGTAACATCCC 1140
QY 1141 GACCATGTTGTTATTAATGATGCTTACGTAGGAGATAGACAGAGCCATGAT 1200
DB 1141 GACCATGTTGTTATTAATGATGCTTACGTAGGAGATAGACAGAGCCATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGAGGTGAGAAACACCATTTGTTGACAAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGAGGTGAGAAACACCATTTGTTGACAAACATGT 1260
QY 1261 GAGGATTCCTTTTGGCTGCTCTATATCTTGAATGCTGCTTCTGCTGAGAGC 1320
DB 1261 GAGGATTCCTTTTGGCTGCTCTATATCTTGAATGCTGCTTCTGCTGAGAGC 1320
QY 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCATCTCAATCCAGCTGCT 1380
DB 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCATCTCAATCCAGCTGCT 1380
QY 1381 ACCATTCCTAGCTATCTGACCAAGGCTCTGTGTTCCACCGGATACCAAGTGGAT 1440
DB 1381 ACCATTCCTAGCTATCTGACCAAGGCTCTGTGTTCCACCGGATACCAAGTGGAT 1440
QY 1441 GATTTGTCAAAGCAGGTGCAATGCTGGAACATTAATGAGGCTTGTGTTGATTTGGC 1500
DB 1441 GATTTGTCAAAGCAGGTGCAATGCTGGAACATTAATGAGGCTTGTGTTGATTTGGC 1500
QY 1501 CCAGAGATTAATCATGATTTCTGAGTCAAGTGA 1533
DB 1501 CCAGAGATTAATCATGATTTCTGAGTCAAGTGA 1533

RESULT 6

US-10-718-952-15

Sequence 15, Application US/10718952

Publication No. US20040128713A1

GENERAL INFORMATION:

APPLICANT: Hiltz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

APPLICANT: Streit, Leon

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-C

QY 1 ATGTTATGCGAATTTTAAGGTAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
DB 1 ATGTTATGCGAATTTTAAGGTAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
QY 61 CAGTCCGTTGTAACCTAGAGAAACCGAAGTGTTCAGAGAACGGAATGGACCTAT 120
DB 61 CAGTCCGTTGTAACCTAGAGAAACCGAAGTGTTCAGAGAACGGAATGGACCTAT 120
QY 121 CAGTGGATTTGCAACCCAAATCCGTCACTACCAATTTAAACCAACCCATGTTTCA 180
DB 121 CAGTGGATTTGCAACCCAAATCCGTCACTACCAATTTAAACCAACCCATGTTTCA 180
QY 181 AAATTTGGGGTGTATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTCAAGGTGT 240
DB 181 AAATTTGGGGTGTATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTCAAGGTGT 240
QY 241 GTTATTTGTAACAGAGAGACATTTTATGAGGCTAACAGAGCAAGATTCAACAGCAAT 300
DB 241 GTTATTTGTAACAGAGAGACATTTTATGAGGCTAACAGAGCAAGATTCAACAGCAAT 300
QY 301 TACTTTGGCTCCCTCACCCAAAGCTCAGACTATTCAGATTTGATCTTCCAGAGAGGAA 360
DB 301 TACTTTGGCTCCCTCACCCAAAGCTCAGACTATTCAGATTTGATCTTCCAGAGAGGAA 360
QY 361 ATCTATGCCCCATTTAAGAGTGTGCTTCAATGTTAATCTCTGACAGCAATGTGTGGG 420
DB 361 ATCTATGCCCCATTTAAGAGTGTGCTTCAATGTTAATCTCTGACAGCAATGTGTGGG 420
QY 421 GGAATGGATATCAGAAATGAACTGCTGATGCTCAAGGCAAGGCAAGGTTGTTGAC 480
DB 421 GGAATGGATATCAGAAATGAACTGCTGATGCTCAAGGCAAGGCAAGGTTGTTGAC 480
QY 481 ATCGATTTGCAAGAGCAATGAGGCTTACATGAGATCCATGTTTCCACTCCCGGAATC 540
DB 481 ATCGATTTGCAAGAGCAATGAGGCTTACATGAGATCCATGTTTCCACTCCCGGAATC 540
QY 541 TACGACCCGGATTTTATGCTGCTGCAACCAAGAGGCTGTCCCAACAGTATTAAGGCG 600
DB 541 TACGACCCGGATTTTATGCTGCTGCAACCAAGAGGCTGTCCCAACAGTATTAAGGCG 600
QY 601 ACAAGCAAGAGCAAGTTTCAAGAAATCATCAAAAGCATCAAGGCTTTAAAGAGCCAC 660
DB 601 ACAAGCAAGAGCAAGTTTCAAGAAATCATCAAAAGCATCAAGGCTTTAAAGAGCCAC 660
QY 661 AAAGTGACAGAGTGTTGCTCTGAGACTGCCAACACAGAGGATGACAAATTTGGTT 720
DB 661 AAAGTGACAGAGTGTTGCTCTGAGACTGCCAACACAGAGGATGACAAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCAATGAGAAATCTCTGGCTGCTGAGACAGAAATGAGCTGAG 780
DB 721 GTAGGCTTTAATGACACCAATGAGAAATCTCTGGCTGCTGAGACAGAAATGAGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGATGCCATTTGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTGATGCCATTTGCTGTGATGAGAAATGTTCTTTCAATTAAT 840

QY 841 GGAAGCCCTCAGAACACTTTTGTATCCAGGCTGATTTGATCTTCCATCCGAGAGAACT 900
DB 841 GGAAGCCCTCAGAACACTTTTGTATCCAGGCTGATTTGATCTTCCATCCGAGAGAACT 900
QY 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTGTGAT 960
DB 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTGTGAT 960
QY 961 TTTCTTTGGGGGCTGTGTATCAAGCCAACTATAGTTAGTTACAACCACTGGGAAAC 1020
DB 961 TTTCTTTGGGGGCTGTGTATCAAGCCAACTATAGTTAGTTACAACCACTGGGAAAC 1020
QY 1021 AATGATGTATGAATCTCTCGGCTCCAAACAACTTCCGCTCCAGAGAAATCTCCAGAGC 1080
DB 1021 AATGATGTATGAATCTCTCGGCTCCAAACAACTTCCGCTCCAGAGAAATCTCCAGAGC 1080
QY 1081 AAGCTTTGACAGATATGTCACACCAATGCCATCCTCTATAGCCTGGTGAACATCCC 1140
DB 1081 AAGCTTTGACAGATATGTCACACCAATGCCATCCTCTATAGCCTGGTGAACATCCC 1140
QY 1141 GACCATTTGTTGTATTTAATATGTGCTTACGTAGGGGATAGCAAGAGCCATGAT 1200
DB 1141 GACCATTTGTTGTATTTAATATGTGCTTACGTAGGGGATAGCAAGAGCCATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATAGGTGAGAAACACCACTTTGTTGACACACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATAGGTGAGAAACACCACTTTGTTGACACACATGT 1260
QY 1261 GAGGATTTCCCTTTTAACTGCTCTATTTATCTTGAATGATCCTTCTTGTGAGTGAAC 1320
DB 1261 GAGGATTTCCCTTTTAACTGCTCTATTTATCTTGAATGATCCTTCTTGTGAGTGAAC 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAAAATGAGGAAAAATTCACCTCATTCACCCAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAAAATGAGGAAAAATTCACCTCATTCACCCAGTTGCT 1380
QY 1381 ACCATTTCTAGCTATCTGACCAAGGCTCCTCTGTGTTCCACCGGTTACACAGTGTGAT 1440
DB 1381 ACCATTTCTAGCTATCTGACCAAGGCTCCTCTGTGTTCCACCGGTTACACAGTGTGAT 1440
QY 1441 GCATTGTCAAGAGCGGTGCAATGCTGAAAAACATATAGGGCTTGTGTGATTTGGCC 1500
DB 1441 GCATTGTCAAGAGCGGTGCAATGCTGAAAAACATATAGGGCTTGTGTGATTTGGCC 1500
QY 1501 CCAGAGATTAACATGATTTCTGAGTACAAAGTGA 1533
DB 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533

RESULT 7

US-10-424-599-12021/c
; Sequence 12021, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12021
; LENGTH: 2582
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1
US-10-424-599-12021

Query Match 64.3%; Score 986; DB 18; Length 2582;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCCCTAAATGTGAATACACCGAGACTGAGATT 60
DB 2454 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCCCTAAATGTGAATACACCGAGACTGAGATT 2395
QY 61 CAGTCCGTGTACACCTACGAAACCAACCGAATTTTTCACGAGAACAGAAATGGCACTAT 120
DB 2394 CAGTCCGTGTACACCTACGAAACCAACCGAATTTTTCACGAGAACAGAAATGGCACTAT 2335
QY 121 CAGTGGATGTCAAAACCCAAATCCGTCACCTACCAATTTAAACCAACACCAATGTTCCA 180
DB 2334 CAGTGGATGTCAAAACCCAAATCCGTCACCTACCAATTTAAACCAACCAACCAATGTTCCA 2275
QY 181 AATTTGGGGGTATGTTGTGGGTTGGGGTGGAAACACGCTCTACCTCAACCGGTGT 240
DB 2274 AATTTGGGGGTATGTTGTGGGTTGGGGTGGAAACACGCTCTACCTCAACCGGTGT 2215
QY 241 GTTATTTCTAACAGAGAGCAATTTTCATAGGCTTACAAAGAACAGATTCACCAAGCAAT 300
DB 2214 GTTATTTCTAACAGAGAGCAATTTTCATAGGCTTACAAAGAACAGATTCACCAAGCAAT 2155
QY 301 TACTTTGGCTCCCTACCCCAAGCTCAGCTATTTCCAGTTGATCTTCCAGGAGAGAA 360
DB 2154 TACTTTGGCTCCCTACCCCAAGCTCAGCTATTTCCAGTTGATCTTCCAGGAGAGAA 2095
QY 361 ATCTATGCCCCATTCAAGAGTCTGCTTCCATAGTGTAAATCTCTGACGACATTTGTTGGG 420
DB 2094 ATCTATGCCCCATTCAAGAGTCTGCTTCCATAGTGTAAATCTCTGACGACATTTGTTGGG 2035
QY 421 GGATGGATATACGAAACATGAACTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 480
DB 2034 GGATGGATATACGAAACATGAACTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1975
QY 481 ATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC 540
DB 1974 ATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC 1915
QY 541 TACGACCCGGAATTTTCAATGCTGCTCCAAACCAAGAGAGCGTCCAAACAGTGTAAAGGCG 600
DB 1914 TACGACCCGGAATTTTCAATGCTGCTCCAAACCAAGAGAGCGTCCAAACAGTGTAAAGGCG 1855
QY 601 ACAAAAGCAAGCAAGTTCAGCAATCATCAAGATCATCAAGCGTTTAAAGAAAGCAC 660
DB 1854 ACAAAAGCAAGCAAGTTCAGCAATCATCAAGATCATCAAGCGTTTAAAGAAAGCAC 1795
QY 661 AAGTGTGACAAAGTGTGTCTGTGTGACTGCAACACAGAGAGTATAGCAATTTGGTT 720
DB 1794 AAGTGTGACAAAGTGTGTCTGTGTGACTGCAACACAGAGAGTATAGCAATTTGGTT 1735
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTGGCTGCTGTGAGACAGAAATGAGGCTGAG 780
DB 1734 GTAGGCTTTAATGACACCATGAGAAATCTTGGCTGCTGTGAGACAGAAATGAGGCTGAG 1675
QY 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATTAAT 840
DB 1674 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATTAAT 1615
QY 841 GGAAGCCCTCAGAACACTTTTGTATCCAGGCTGATTTGATCTTCCATCCGAGAGAACT 900
DB 1614 GGAAGCCCTCAGAACACTTTTGTATCCAGGCTGATTTGATCTTCCATCCGAGAGAACT 1555
QY 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTGTGAT 960
DB 1554 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTGTGAT 1495
QY 961 TTTCTTTGGGGGCTGTGTATCAAGCCAACTATAGTTAGTTACAACCACTGGGAAAC 1020
DB 1494 TTTCTTTGGGGGCTGTGTATCAAGCCAACTATAGTTAGTTACAACCACTGGGAAAC 1435
QY 1021 AATGATGTATGAATCTCTCGGCTCCAAACAACTTCCGCTCCAGAGAAATCTCCAGAGC 1080

Db 1434 AATGATGATGAATCTCGGCTCCACAAACCTTCGCTCCAGAAATCTCCAAAGAGC 1375
QY 1081 AACGTTGAGATATGAGTCAACAGCAATGCATCTCTATAGAGCTGGAGATCC 1139
Db 1374 AACGTTGAGATATGAGTCAACAGCAATGCATCTCTATAGAGCTGGAGATCC 1316

RESULT 8
US-10-424-599-22663

/ Sequence 22663, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 22663
/ LENGTH: 634
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_120467C.1
US-10-424-599-22663

Query Match 20.2%; Score 310; DB 18; Length 634;
Best Local Similarity 100.0%; Pred. No. 8.3e-156;

Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 705 GATAGCAATTTGGTTGAGCCTTAATGACACCATGAGAAATCTCTGGCTGCTGGA 764
Db 4 GATAGCAATTTGGTTGAGCCTTAATGACACCATGAGAAATCTCTGGCTGCTGGA 63
QY 765 CAGAAATGAGGCTGATTTCTCCCTGATATGCGATTCGCTGATGAGAA 824
Db 64 CAGAAATGAGGCTGATTTCTCCCTGATATGCGATTCGCTGATGAGAA 123
QY 825 TGTTCCTTCATTAATGGAAGCCTCAGAACATTTTGTACAGAGGCTGATGATCTGC 884
Db 124 TGTTCCTTCATTAATGGAAGCCTCAGAACATTTTGTACAGAGGCTGATGATCTGC 183
QY 885 CATCGGAGAACATTTTATGAGATGATGATCAAGATGCTCAGACCAAAATGAA 944
Db 184 CATCGGAGAACATTTTATGAGATGATGATCAAGATGCTCAGACCAAAATGAA 243
QY 945 ATCTGTTGTTGATTTCTTGTGGGGGCTGATCAAGCCAACTATATGTTAGTTA 1004
Db 244 ATCTGTTGTTGATTTCTTGTGGGGGCTGATCAAGCCAACTATATGTTAGTTA 303
QY 1005 CAACCATCTG 1014
Db 304 CAACCATCTG 313

RESULT 9
US-10-424-599-70162

/ Sequence 70162, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599

/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684

/ SEQ ID NO 70162

/ LENGTH: 844

/ TYPE: DNA

/ ORGANISM: Glycine max

/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_34368C.1
US-10-424-599-70162Query Match 12.1%; Score 185; DB 18; Length 844;
Best Local Similarity 100.0%; Pred. No. 2e-88;

Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1167 GCCTTACGTAGGGGATATGACAGAGCCATGATGATGATCACTTCAGAGATATTCATGGG 1226
Db 9 GCCTTACGTAGGGGATATGACAGAGCCATGATGATGATCACTTCAGAGATATTCATGGG 68
QY 1227 TGGAAAGAACACATTTGTTGACACACATGTGAGGATTCCTTTAGCTGCTCTAT 1286
Db 69 TGGAAAGAACACATTTGTTGACACACATGTGAGGATTCCTTTAGCTGCTCTAT 128
QY 1287 TATCTTGACCTTGTGCTCTTGTGAGCTGAGCACTGAAATCCAGTTAAAGCTGAA 1346
Db 129 TATCTTGACCTTGTGCTCTTGTGAGCTGAGCACTGAAATCCAGTTAAAGCTGAA 188
QY 1347 TGAGG 1351
Db 189 TGAGG 193

RESULT 10
US-10-424-599-116597

/ Sequence 116597, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 116597
/ LENGTH: 661
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)-(661)
/ OTHER INFORMATION: unsure at all n locations
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_76298C.1
US-10-424-599-116597

Query Match 9.5%; Score 146; DB 18; Length 661;
Best Local Similarity 100.0%; Pred. No. 2e-67;

Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGAAATGTTCTCTTATTAATGAGCCCTCAGAACATTTTGTATCAGAGGCTGATT 876
Db 56 ATGAAATGTTCTCTTATTAATGAGCCCTCAGAACATTTTGTATCAGAGGCTGATT 115
QY 877 GATCTTGCATCGCAGAGAACATTTGATGTGTGAGATGATCTCAAGATGTGCAACC 936
Db 116 GATCTTGCATCGCAGAGAACATTTGATGTGTGAGATGATCTCAAGATGTGCAACC 175
QY 937 AAAATGAATCTGTGTGTTGATT 962
Db 176 AAAATGAATCTGTGTGTTGATT 201

```
RESULT 11
US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-5

Query Match
Best Local Similarity 100.0%; Score 146; DB 14; Length 1533;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGAAATGTTCTTTCATTATGAGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
DB 817 ATGAAATGTTCTTTCATTATGAGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
QY 877 GATCTTGCCATGCGAGGAGAACCTTTGATGAGATGACTTCAAGAGTGTGACACC 936
DB 877 GATCTTGCCATGCGAGGAGAACCTTTGATGAGATGACTTCAAGAGTGTGACACC 936
QY 937 AAAATGAATCTGTGTGTTGATT 962
DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 12
US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-9

Query Match
Best Local Similarity 100.0%; Score 146; DB 14; Length 1533;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGAAATGTTCTTTCATTATGAGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
DB 817 ATGAAATGTTCTTTCATTATGAGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
QY 877 GATCTTGCCATGCGAGGAGAACCTTTGATGAGATGACTTCAAGAGTGTGACACC 936
DB 877 GATCTTGCCATGCGAGGAGAACCTTTGATGAGATGACTTCAAGAGTGTGACACC 936
QY 937 AAAATGAATCTGTGTGTTGATT 962
DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 13
US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-5

Query Match
Best Local Similarity 100.0%; Score 146; DB 19; Length 1533;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGAAATGTTCTTTCATTATGAGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
DB 817 ATGAAATGTTCTTTCATTATGAGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
QY 877 GATCTTGCCATGCGAGGAGAACCTTTGATGAGATGACTTCAAGAGTGTGACACC 936
DB 877 GATCTTGCCATGCGAGGAGAACCTTTGATGAGATGACTTCAAGAGTGTGACACC 936
QY 937 AAAATGAATCTGTGTGTTGATT 962
DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 14
US-10-718-952-9
; Sequence 9, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:47:55 ; Search time 92 Seconds
(without alignments)
2143.998 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 510
Sequence: 1 MFIEFKVESNVKXTEREI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq16Dec04:*

- 1: geneseqp19808:*
- 2: geneseqp19908:*
- 3: geneseqp20008:*
- 4: geneseqp20018:*
- 5: geneseqp20028:*
- 6: geneseqp20038:*
- 7: geneseqp20038:*
- 8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	510	8	ADQ14501 Mutant so
2	510	100.0	510	8	AD882002 Soybean m
3	423	82.9	510	8	ADQ14503 Mutant so
4	423	82.9	510	8	ADQ14505 Wild type
5	423	82.9	510	8	AD882004 Soybean m
6	423	82.9	510	8	AD882006 Soybean m
7	236	46.3	510	2	AAW79740 Soybean w
8	236	46.3	510	8	ADQ14499 Mutant so
9	236	46.3	510	8	ADQ14491 Wild type
10	236	46.3	510	8	AD882000 Soybean m
11	236	46.3	510	8	AD881994 Soybean m
12	220	43.1	510	2	AAW79741 Soybean m
13	220	43.1	510	8	ADQ14495 Mutant so
14	220	43.1	510	8	AD881998 Soybean m
15	76	14.9	510	2	AAW79782 Maize myo
16	76	14.9	510	2	AAW96259 Phytate P
17	76	14.9	512	8	ADP43919 Porterei
18	76	14.9	536	2	ADP43919 Porterei
19	67	13.1	380	3	AAZ24477 Nicotiana
20	67	13.1	380	3	AAZ24477 Nicotiana
21	67	13.1	446	3	AAZ32503 Arabidops
22	67	13.1	446	3	AAZ32502 Arabidops
23	67	13.1	505	8	AAZ32502 Arabidops
24	67	13.1	510	3	AAZ32502 Arabidops
25	67	13.1	510	4	AAZ32502 Arabidops

26	67	13.1	510	8	ADN73525	Adh73525 Thale cre
27	67	13.1	510	8	ADP43920	Adp43920 Rice l-my
28	67	13.1	511	3	AAZ09861	AAZ09861 Arabidops
29	67	13.1	511	3	AAZ32501	AAZ32501 Arabidops
30	67	13.1	534	3	AAZ09860	AAZ09860 Arabidops
31	67	13.1	581	3	AAZ05074	AAZ05074 Arabidops
32	67	13.1	645	3	AAZ05073	AAZ05073 Arabidops
33	44	8.6	394	5	ABP51436	ABP51436 Human MDD
34	32	6.3	98	3	AAZ44825	AAZ44825 Zea mays
35	29	5.7	565	4	ABZ65497	ABZ65497 Drosophila
36	20	3.9	45	3	AAZ05076	AAZ05076 Arabidops
37	20	3.9	78	5	ABP07716	ABP07716 Human ORF
38	20	3.9	113	3	AAZ12422	AAZ12422 Zea mays
39	20	3.9	128	3	AAZ32705	AAZ32705 Zea mays
40	20	3.9	130	3	AAZ32704	AAZ32704 Zea mays
41	20	3.9	138	3	AAZ19004	AAZ19004 Zea mays
42	20	3.9	153	3	AAZ32703	AAZ32703 Zea mays
43	20	3.9	174	3	AAZ19003	AAZ19003 Zea mays
44	19	3.7	251	3	AAZ52875	AAZ52875 Arabidops
45	19	3.7	316	3	AAZ52874	AAZ52874 Arabidops

ALIGNMENTS

RESULT 1	
ID ADQ14501	standard; protein; 510 AA.
XX	
AC ADQ14501;	
XX	
DT 23-SEP-2004	(first entry)
XX	
DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #3.	
XX	
KW soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;	
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;	
KW mutant; mutlein.	
XX	
OS Glycine max.	
OS Synthetic.	
XX	
PN US2004128713-A1.	
XX	
PD 01-JUL-2004.	
XX	
PF 21-NOV-2003; 2003US-00718952.	
XX	
PR 08-APR-1997; 97US-00835751.	
PR 07-APR-1998; 98WO-US006822.	
PR 26-APR-1999; 99US-0029315.	
PR 11-MAR-2002; 2002US-00025003.	
XX	
PA (HITZ/) HITZ W D.	
PA (SEBA/) SEBASTIAN S A.	
PA (GRAC/) GRACE D J.	
PA (STRE/) STREIT L G.	
XX	
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
XX	
DR MPI: 2004-53135/51.	
XX	
N-PSDB; ADQ14500.	
XX	
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,	
PT useful for producing plants with decreased raffinose, stachyose, and	
PT phytic acid and increased sucrose, leading to valuable and useful soybean	
PT products.	
XX	
PS Claim 9, SEQ ID NO 12; 48pp; English.	
XX	
CC The invention relates to a nucleic acid fragment encoding a soybean myo-	
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate	
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-	

CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MFIENFKVESPNNVKTETETIOSVYNYETTELVHENRNGTYQWIVKPKSVNYQFKTNTHVP 60
DB 1 MFIENFKVESPNNVKTETETIOSVYNYETTELVHENRNGTYQWIVKPKSVNYQFKTNTHVP 60
QY 61 KLGVNLVGMGNGNSTLTGVTIANREDISWATKDKIQQANFGSLTQASAIRVGSFQGBE 120
DB 61 KLGVNLVGMGNGNSTLTGVTIANREDISWATKDKIQQANFGSLTQASAIRVGSFQGBE 120
QY 121 IYAPFKSLPLPVNPPDIVFGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESWVPLPGI 180
DB 121 IYAPFKSLPLPVNPPDIVFGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESWVPLPGI 180
QY 181 YDPDFIAANOEBRANNVIKGTQEOVOQIIKDKIAFKKATKDKVYVMTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVIKGTQEOVOQIIKDKIAFKKATKDKVYVMTANTERYSNLV 240
QY 241 VGLNNTMENTLAAVRNEAEISPTLTVAICMENVPIFNGSPONTFPGILDIAIANT 300
DB 241 VGLNNTMENTLAAVRNEAEISPTLTVAICMENVPIFNGSPONTFPGILDIAIANT 300
QY 301 LIGGDDFKSGQTKMSVLDPLVGAIGIKPTISVYNHLGNDGNMISAPOFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLDPLVGAIGIKPTISVYNHLGNDGNMISAPOFRSKEISKS 360
QY 361 NVVDMVNSNALTYEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSBIFMGKXTIVLHNTC 420
DB 361 NVVDMVNSNALTYEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSBIFMGKXTIVLHNTC 420
QY 421 EDSLAAPIILDVLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKPLVPPGTPVYN 480
DB 421 EDSLAAPIILDVLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKPLVPPGTPVYN 480
QY 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510
```

RESULT 2
ADS82002
ID ADS82002 standard; protein; 510 AA.

AC ADS82002;

DT 18-NOV-2004 (first entry)

XX Soybean myo-inositol 1-phosphate synthase mutant #2.

KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
KM phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
KM flaculence; mutant.

XX

OS Glycine max; line 29010CP01.
OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 87 /note= "Wild-type Gly substituted by Asp"

XX US2003074685-A1.

XX 17-APR-2003.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX Hitz WD, Sebastian SA;

XX WPI, 2004-639957/62.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.

XX Claim 9; SEQ ID NO 12; 34pp; English.

The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreased capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flaculence
CC when consumed by humans. The present sequence represents a mutant myo-
CC inositol 1-phosphate synthase.

SQ Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFIENFKVESPNNVKTETETIOSVYNYETTELVHENRNGTYQWIVKPKSVNYQFKTNTHVP 60

DB 1 MFIENFKVESPNNVKTETETIOSVYNYETTELVHENRNGTYQWIVKPKSVNYQFKTNTHVP 60

QY 61 KLGVNLVGMGNGNSTLTGVTIANREDISWATKDKIQQANFGSLTQASAIRVGSFQGBE 120

|||||

DB 61 KLGWMLVGKGNNGSLTNGVIANREDISWATKDKIQANVFSGSLTQASAIRVGSFQGE 120
QY 121 ITAPFKSLPKNPPDIYVGGWDISNMNADAMARAKVPDILQQLRPYEMSWPLPGI 180
DB 121 ITAPFKSLPKNPPDIYVGGWDISNMNADAMARAKVPDILQQLRPYEMSWPLPGI 180
QY 181 YPDPFLAANQBERANNVIGTKQBOVOQIIKDIAKFKETKVDKVVVMTANTERYSNLY 240
DB 181 YPDPFLAANQBERANNVIGTKQBOVOQIIKDIAKFKETKVDKVVVMTANTERYSNLY 240
QY 241 VGLNDTMENLLAADVRENEAISPSTLYAACWENVPFINGSPONTFVPGILDLAIARNT 300
DB 241 VGLNDTMENLLAADVRENEAISPSTLYAACWENVPFINGSPONTFVPGILDLAIARNT 300
QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKRSTISVSYNHLGNNDGMISAPQTRSKSISKS 360
DB 301 LIGGDDFKSGQTKMSVLVDPLVAGIKRSTISVSYNHLGNNDGMISAPQTRSKSISKS 360
QY 361 NVVDDMVNSNALLYEPGEHPDHVVVVKYVPYVGDSCRADDEYTSIEIPMGKNTIVLHNTC 420
DB 361 NVVDDMVNSNALLYEPGEHPDHVVVVKYVPYVGDSCRADDEYTSIEIPMGKNTIVLHNTC 420
QY 421 EDSLAAPIILDLVLAELSTRIQFKAENEKGFHSPVATILSYLTAKAPLVPPTPVN 480
DB 421 EDSLAAPIILDLVLAELSTRIQFKAENEKGFHSPVATILSYLTAKAPLVPPTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
RESULT 3
ADQ14503
ID ADQ14503 standard; protein; 510 AA.
AC ADQ14503;
XX
DT 23-SEP-2004 (first entry)
DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #4.
XX
KM Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
XX
OS Glycine max.
OS Synthetic.
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
PF 21-NOV-2003; 2003US-00718952.
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1996; 98WO-US006822.
PR 26-APR-1999; 99US-00293315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
DR MPI: 2004-53335/51.
DR N-PSDB; ADQ14502.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX
PS Example 8; SEQ ID NO 14; 48bp; English.
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.
CC
SQ Sequence 510 AA;
Query Match 82.9%; Score 423; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 ISWATKDKIQANVFSGSLTQASAIRVGSFQGEIYAPFKSLPKNPPDIYVGGWDISNM 147
DB 88 ISWATKDKIQANVFSGSLTQASAIRVGSFQGEIYAPFKSLPKNPPDIYVGGWDISNM 147
QY 148 NIADMARAKVPDILQQLRPYEMSWPLPGIYDPDFIAANQBERANNVIGTKQBOVO 207
DB 148 NIADMARAKVPDILQQLRPYEMSWPLPGIYDPDFIAANQBERANNVIGTKQBOVO 207
QY 208 QIITKIKAFKATKVDKVVVMTANTERYSNLVGLNDTMENLLAADVRENEAISPSTLY 267
DB 208 QIITKIKAFKATKVDKVVVMTANTERYSNLVGLNDTMENLLAADVRENEAISPSTLY 267
QY 268 AIACWENVPFINGSPONTFVPGILDLAIARNTLIGGDDFKSGQTKMSVLVDPLVAGI 327
DB 268 AIACWENVPFINGSPONTFVPGILDLAIARNTLIGGDDFKSGQTKMSVLVDPLVAGI 327
QY 328 KPTSIYSYNHLGNNDGMISAPQTRSKSISKNVDDMVNSNALLYEPGEHPDHVVVIX 387
DB 328 KPTSIYSYNHLGNNDGMISAPQTRSKSISKNVDDMVNSNALLYEPGEHPDHVVVIX 387
QY 388 YVPYVGDSCRADDEYTSIEIPMGKNTIVLHNTCEBSLLAAPILDLVLAELSTRIQFKA 447
DB 388 YVPYVGDSCRADDEYTSIEIPMGKNTIVLHNTCEBSLLAAPILDLVLAELSTRIQFKA 447
QY 448 ENEGKGFHSPVATILSYLTAKAPLVPPTPVNALSOKRAMLENIMRACVGLAPENNMIL 507
DB 448 ENEGKGFHSPVATILSYLTAKAPLVPPTPVNALSOKRAMLENIMRACVGLAPENNMIL 507
QY 508 EYK 510
DB 508 EYK 510
RESULT 4
ADQ14505
ID ADQ14505 standard; protein; 510 AA.
AC ADQ14505;
XX
DT 23-SEP-2004 (first entry)
DE Wild type soybean myo-inositol 1-phosphate synthase polypeptide #2.
XX
KM Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.


```

XX Glycine max.
OS US2004128713-A1.
PN 01-JUL-2004.
PF 21-NOV-2003; 2003US-00718952.
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-0029315.
PR 11-MAR-2002; 2002US-00025003.
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI: 2004-533135/51.
XX N-PSDB; ADQ14504.
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX Claim 3; SEQ ID NO 16; 48pp; English.
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents a wild type soybean myo-
XX inositol 1-phosphate synthase polypeptide of the invention.
SQ Sequence 510 AA;
Query Match 82.9%; Score 423; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 328 KPTSIIVSYNHLGNDGNMISAPQTFRSKEISKSNVVDWVNSNALYEPGEHPHVVIK 387
QY 388 YVPVVGDSKRAMDEXTSEIFMGSKNTIYLAHNTCEDSLAAPIIDLVYLAELSTRIOFKA 447
DB 388 YVPVVGDSKRAMDEXTSEIFMGSKNTIYLAHNTCEDSLAAPIIDLVYLAELSTRIOFKA 447
QY 448 ENEGKFHSHFHVATILSYLTKAPLVPPEPTPVVNALSKORAMLENIMRACVGLAENNMIL 507
DB 448 ENEGKFHSHFHVATILSYLTKAPLVPPEPTPVVNALSKORAMLENIMRACVGLAENNMIL 507
QY 508 EYK 510
DB 508 EYK 510
RESULT 5
ADSS82004
ID ADSS82004 standard; protein; 510 AA.
XX
AC ADSS82004;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase wild-type 4.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 290180P03.
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PI Hitz WD, Sebastian SA;
XX
DR WPI: 2004-639957/62.
XX N-PSDB; ADS82003.
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.
XX
PS Example 8; SEQ ID NO 14; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreased capacity for the synthesis of myo-
XX inositol 1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g), provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived
XX from seeds of a soybean plant (homozygous for one or more gene encoding a

```

CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC a heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.

Sequence 510 AA;

Query Match 82.9%; Score 423; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ISMATKDKIQOANFGSLTQASAIRVSGFQGEIYAPKSLIPMNPPDIVFGWDISNM 147
 DB 88 ISMATKDKIQOANFGSLTQASAIRVSGFQGEIYAPKSLIPMNPPDIVFGWDISNM 147
 QY 148 NLADAMAAKAFDIDLQKQLRPYMESWVPLPGIYDPDPIAANOBERANNAVIGTKQEOVQ 207
 DB 148 NLADAMAAKAFDIDLQKQLRPYMESWVPLPGIYDPDPIAANOBERANNAVIGTKQEOVQ 207
 QY 208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGINDTMENLLAADVNEAEISPSSTLY 267
 DB 208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGINDTMENLLAADVNEAEISPSSTLY 267
 QY 268 AIACMVENVPIFNGSPONTFVPGGLDLAIARNTLIGDDDFKSGQTKMSVLDVFLVAGGI 327
 DB 268 AIACMVENVPIFNGSPONTFVPGGLDLAIARNTLIGDDDFKSGQTKMSVLDVFLVAGGI 327
 QY 328 KPTSIIVSNHLGNNDGNNLSAPQTFRSKEISKSNNVDDMVNSNAILYEPGHPDHVVVIX 387
 DB 328 KPTSIIVSNHLGNNDGNNLSAPQTFRSKEISKSNNVDDMVNSNAILYEPGHPDHVVVIX 387
 QY 388 YVPYVGDSTRANDERTSEIFMGCKNTIVLANTCEDSLLAAPILLDLVLIAELSTRIOFKA 447
 DB 388 YVPYVGDSTRANDERTSEIFMGCKNTIVLANTCEDSLLAAPILLDLVLIAELSTRIOFKA 447
 QY 448 ENEGKFHSHFPAATLISVYTKAPLVPGRPVVNNALSKORAMENIMRACVGLAPENNMIL 507
 DB 448 ENEGKFHSHFPAATLISVYTKAPLVPGRPVVNNALSKORAMENIMRACVGLAPENNMIL 507
 QY 508 EYK 510
 DB 508 EYK 510

RESULT 6
 ADS82006
 ID ADS82006 standard; protein; 510 AA.
 AC ADS82006;
 DT 18-NOV-2004 (first entry)
 DE Soybean myo-inositol 1-phosphate synthase wild-type 2.
 KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
 KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 OS Glycine max; cultivar Wye.
 PN US2003074685-A1.
 PD 17-APR-2003.
 PF 11-MAR-2002; 2002US-00025003.
 PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.
 XX (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 XX Hitz WD, Sebastian SA;
 XX MPl. 2004-639957/62.
 DR N-PSDB; ADS82005.
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 Claim 3, SEQ ID NO 16; 34pp; English.

The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreased capacity for the synthesis of myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g), provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC a heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.

Sequence 510 AA;

Query Match 82.9%; Score 423; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ISMATKDKIQOANFGSLTQASAIRVSGFQGEIYAPKSLIPMNPPDIVFGWDISNM 147
 DB 88 ISMATKDKIQOANFGSLTQASAIRVSGFQGEIYAPKSLIPMNPPDIVFGWDISNM 147
 QY 148 NLADAMAAKAFDIDLQKQLRPYMESWVPLPGIYDPDPIAANOBERANNAVIGTKQEOVQ 207
 DB 148 NLADAMAAKAFDIDLQKQLRPYMESWVPLPGIYDPDPIAANOBERANNAVIGTKQEOVQ 207
 QY 208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGINDTMENLLAADVNEAEISPSSTLY 267
 DB 208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGINDTMENLLAADVNEAEISPSSTLY 267
 QY 268 AIACMVENVPIFNGSPONTFVPGGLDLAIARNTLIGDDDFKSGQTKMSVLDVFLVAGGI 327
 DB 268 AIACMVENVPIFNGSPONTFVPGGLDLAIARNTLIGDDDFKSGQTKMSVLDVFLVAGGI 327
 QY 328 KPTSIIVSNHLGNNDGNNLSAPQTFRSKEISKSNNVDDMVNSNAILYEPGHPDHVVVIX 387
 DB 328 KPTSIIVSNHLGNNDGNNLSAPQTFRSKEISKSNNVDDMVNSNAILYEPGHPDHVVVIX 387

QY 388 YVPYVGDSSGRANDEYTSSEIFMGSKNTIVLHANTCEDSLAAPILLDLVLAELSTRIOEKA 447
DB 388 YVPYVGDSSGRANDEYTSSEIFMGSKNTIVLHANTCEDSLAAPILLDLVLAELSTRIOEKA 447
QY 448 EHEGKFSHFHPVATILTSYLTKAPLVPPGPPVYNALSKORAMENIMRACVGLAPNNMIL 507
DB 448 EHEGKFSHFHPVATILTSYLTKAPLVPPGPPVYNALSKORAMENIMRACVGLAPNNMIL 507
QY 508 EYK 510
DB 508 EYK 510

RESULT 7
AAW79740
ID AAW79740 standard; protein; 510 AA.

AC AAW79740;
DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)

DE Soybean wild-type myo-inositol 1-phosphate synthase.

KW Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW phytic acid.

OS Glycine max; line LR13.

PN W09845448-A1.

PD 15-OCT-1998.

PF 07-APR-1998; 98WO-US006822.

PR 08-APR-1997; 97US-00835751.

(DUPO) DU PONT DE NEMOURS & CO E I.

Hitz WD, Sebastian SA;

WPI: 1998-568353/48.

N-PSDB; AAV62440.

PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.

Example 5; Page 45-47; 63pp; English.

PS This is the amino acid sequence of soybean myo-inositol 1-phosphate
XX synthase (MI 1-PS) deduced from the coding region of an isolated cDNA
CC clone (see AAV62440). MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. A mutant MI 1-PS (see AAW79741) has been
CC identified in soybean line LR13, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC in the LR13 gene sequence that resulted in a K396N substitution in the
CC mutant protein. The mutation results in a seed phenotype of very low
CC raffinose saccharide sugars, very high sucrose and low phytic acid. The
CC mutated nucleic acid is used to alter the raffinose saccharide, sucrose,
CC phytic acid and inorganic phosphate content of soybean seeds, leading to
CC useful soybean products, e.g. a seed phytic acid content of less than 17
CC ug/g, a seed content of raffinose and stachyose combined of less than 17
CC 14.5 ug/g, and a seed sucrose content greater than 200 ug/g. (Updated on
CC 17-OCT-2003 to standardise OS field)

XX Sequence 510 AA;

Query Match 46.3%; Score 236; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.5e-227;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPDFIAAOEBRANNVIKGTQEOVOQIHKIKAFKATKVDKVVVLTWANTER 235

DB 176 PLPGIYDPDFIAAOEBRANNVIKGTQEOVOQIHKIKAFKATKVDKVVVLTWANTER 235
QY 236 YSNLVGLNDTMENILAAVDNREAEISPSLYALACVMENVPFINGSPONTFVGLIDLA 295
DB 236 YSNLVGLNDTMENILAAVDNREAEISPSLYALACVMENVPFINGSPONTFVGLIDLA 295
QY 296 IARNTLIGDDPFKSGQTKMSVLVDFLVGAGIKPSTIVSYNHLGNDGMNISAPQTERSK 355
DB 296 IARNTLIGDDPFKSGQTKMSVLVDFLVGAGIKPSTIVSYNHLGNDGMNISAPQTERSK 355
QY 356 EISKSNVVDMDVNSNALIYEPGEHPDHVVVYKYVPYVGDSSGRANDEYTSSEIFMGK 411
DB 356 EISKSNVVDMDVNSNALIYEPGEHPDHVVVYKYVPYVGDSSGRANDEYTSSEIFMGK 411

RESULT 8
ADQ14499
ID ADQ14499 standard; protein; 510 AA.

AC ADQ14499;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #2.

XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;

KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;

KW mutant; mutein.

OS Glycine max.

OS Synthetic.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.

PR 11-MAR-2002; 2002US-00025003.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

Hitz WD, Sebastian SA, Grace DJ, Streit LG;

WPI: 2004-533135/51.

N-PSDB; ADQ14498.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Example 8; SEQ ID NO 10; 48pp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having

CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

XX
SQ Sequence 510 AA;

Query Match 46.3%; Score 236; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.5e-227;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPFIANOEERANNVTKGTROBOVOQIIKDIAKFEATKVDKVVVLTWNTNR 235
DB 176 PLPGIYDPFIANOEERANNVTKGTROBOVOQIIKDIAKFEATKVDKVVVLTWNTNR 235
QY 236 YSNLVVGLNDTMENTLAAVDRNEAIEISPTLYAICWENVPFINGSPONTFVPGIIDA 295
DB 236 YSNLVVGLNDTMENTLAAVDRNEAIEISPTLYAICWENVPFINGSPONTFVPGIIDA 295
QY 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
DB 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
QY 356 EISKSNVVDDMVNSNAILYEPGEHPDHVVVTKVPYVGDGSKRAMDEYTSIEIFMGCK 411
DB 356 EISKSNVVDDMVNSNAILYEPGEHPDHVVVTKVPYVGDGSKRAMDEYTSIEIFMGCK 411

RESULT 9
ADQ14491
ID ADQ14491 standard; protein; 510 AA.

XX ADQ14491;

DT 23-SEP-2004 (first entry)

XX Wild type soybean myo-inositol 1-phosphate synthase polypeptide #1.

XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.

XX Glycine max.

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 26-APR-1999; 99US-00299315.

XX 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX (GRAC/) GRACE D J.

XX (STRE/) STREIT L G.

XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI; 2004-533135/51.

XX N-PSDB; ADQ14490.

XX Claim 3; SEQ ID NO 2; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a wild type soybean myo-
CC inositol 1-phosphate synthase polypeptide of the invention.

XX
SQ Sequence 510 AA;

Query Match 46.3%; Score 236; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.5e-227;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPFIANOEERANNVTKGTROBOVOQIIKDIAKFEATKVDKVVVLTWNTNR 235
DB 176 PLPGIYDPFIANOEERANNVTKGTROBOVOQIIKDIAKFEATKVDKVVVLTWNTNR 235
QY 236 YSNLVVGLNDTMENTLAAVDRNEAIEISPTLYAICWENVPFINGSPONTFVPGIIDA 295
DB 236 YSNLVVGLNDTMENTLAAVDRNEAIEISPTLYAICWENVPFINGSPONTFVPGIIDA 295
QY 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
DB 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
QY 356 EISKSNVVDDMVNSNAILYEPGEHPDHVVVTKVPYVGDGSKRAMDEYTSIEIFMGCK 411
DB 356 EISKSNVVDDMVNSNAILYEPGEHPDHVVVTKVPYVGDGSKRAMDEYTSIEIFMGCK 411

RESULT 10
ADS82000
ID ADS82000 standard; protein; 510 AA.

XX ADS82000;

DT 18-NOV-2004 (first entry)

XX Soybean myo-inositol 1-phosphate synthase wild-type 3.

XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;

KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

XX Glycine max; line 29004UP01.

XX US2003074685-A1.

XX 17-APR-2003.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX Hitz WD, Sebastian SA;

XX WPI; 2004-639957/62.

XX N-PSDB; ADS81999.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Example 8; SEQ ID NO 10; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a wild-type myo-inositol 1-phosphate synthase.

XX Sequence 510 AA;

XX Query Match 46.3%; Score 236; DB 8; Length 510;

XX Best Local Similarity 100.0%; Pred. No. 4.5e-227; Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 176 PLPGIYDPPFIANAQBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVLTANTER 235
XX 176 PLPGIYDPPFIANAQBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVLTANTER 235
XX 236 YSNLVVGLNDTMENTLLAAVDNRNEAISPSTLYAIAACWENVPFINGSPQNTFVGLIDLA 295
XX 236 YSNLVVGLNDTMENTLLAAVDNRNEAISPSTLYAIAACWENVPFINGSPQNTFVGLIDLA 295
XX 236 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGNTSAPQTFPSK 355
XX 236 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGNTSAPQTFPSK 355
XX 356 EISKSNVVDMDMNSAALTEPGEHPDHVVVYKVPVYGSKRAMDEYTSIFMGSK 411
XX 356 EISKSNVVDMDMNSAALTEPGEHPDHVVVYKVPVYGSKRAMDEYTSIFMGSK 411

XX RESULT 11
XX ADS81994
XX ID ADS81994 standard; protein; 510 AA.

XX ADS81994;

XX 18-NOV-2004 (first entry)

XX Soybean myo-inositol 1-phosphate synthase wild-type 1.

XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid; raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

XX Glycine max; cultivar Wye.

XX US2003074685-A1.

XX 17-APR-2003.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX Hitz WD, Sebastian SA;

XX WPI, 2004-639957/62.

XX N-PsDB; ADS81993.

XX Claim 3; SEQ ID NO 2; 34pp; English.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a wild-type myo-inositol 1-phosphate synthase.

XX Sequence 510 AA;

XX Query Match 46.3%; Score 236; DB 8; Length 510;

XX Best Local Similarity 100.0%; Pred. No. 4.5e-227; Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 176 PLPGIYDPPFIANAQBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVLTANTER 235
XX 176 PLPGIYDPPFIANAQBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVLTANTER 235

XX 236 YSNLVVGLNDTMENTLLAAVDNRNEAISPSTLYAIAACWENVPFINGSPQNTFVGLIDLA 295
XX 236 YSNLVVGLNDTMENTLLAAVDNRNEAISPSTLYAIAACWENVPFINGSPQNTFVGLIDLA 295

XX 236 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGNTSAPQTFPSK 355
XX 236 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNDGNTSAPQTFPSK 355

```

Db      296 IARNLTIGGDDFKSGOTKKKSVLVDPLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
Qy      356 EISKSNVDDMNNSNAILYEPGEHPDHVVVVIKVPYVGSKRAMDYTSIEIFMGK 411
      |||
      356 EISKSNVDDMNNSNAILYEPGEHPDHVVVVIKVPYVGSKRAMDEYTSIEIFMGK 411

RESULT 12
AAM79741
ID      AAM79741 standard; protein; 510 AA.
XX
AC      AAM79741;
XX
DT      17-OCT-2003 (revised)
DT      02-FEB-1999 (first entry)
XX
DE      Soybean mutant myo-inositol 1-phosphate synthase.
XX
KW      Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW      phytic acid.
XX
OS      Glycine max; line LR33.
XX
PN      WO9845448-A1.
XX
PD      15-OCT-1998.
XX
PF      07-APR-1998; 98WO-US006822.
XX
PR      08-APR-1997; 97US-00835751.
XX
PA      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI      Hitz WD, Sebastian SA;
XX
DR      WPI; 1998-568353/48.
DR      N-PSDB; AAV62443.
XX
PT      Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT      for generating plants with altered levels of e.g. raffinose, stachyose,
PT      phytic acid, etc.
XX
PS      Example 5; Page 49-51; 63pp; English.
XX
CC      This is the amino acid sequence of a mutant soybean myo-inositol 1-
CC      phosphate synthase (MI 1-PS) deduced from the coding region of an
CC      isolated cDNA clone (see AAV62443). MI 1-PS is involved in glucose
CC      metabolism to phytic acid, raffinose and stachyose. The MI 1-PS was
CC      identified in soybean line LR33, a mutagenised line of low raffinose
CC      saccharide phenotype. Sequencing revealed a single base change mutation
CC      in the LR33 gene sequence that resulted in a K396N substitution in the
CC      mutant protein compared to wild-type MI 1-PS (see AAM79740). The mutation
CC      results in a seed phenotype of very low raffinose saccharide sugars, very
CC      high sucrose and low phytic acid. The mutated nucleic acid is used to
CC      alter the raffinose saccharide, sucrose, phytic acid and inorganic
CC      phosphate content of soybean seeds, leading to useful soybean products,
CC      e.g. a seed phytic acid content of less than 17 ug/g, a seed content of
CC      raffinose and stachyose combined of less than 14.5 ug/g, a seed content of
CC      sucrose content greater than 200 ug/g. (Updated on 17-OCT-2003 to
CC      standardise OS field)
XX
SQ      Sequence 510 AA;

Query Match      43.1%; Score 220; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.7e-211;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      176 PIRGTYDPPFIANQGERANNVIKGTROQVOQIIDIAFKFATVDKVVVLTMTANTER 235
      |||
      176 PIRGTYDPPFIANQGERANNVIKGTROQVOQIIDIAFKFATVDKVVVLTMTANTER 235
Db      176 PIRGTYDPPFIANQGERANNVIKGTROQVOQIIDIAFKFATVDKVVVLTMTANTER 235
      |||
Qy      236 YSNLVGLNDTMENTLAAVDNRNEAETSPETLYAACMENVEPFINGSPONTFVPGILIDA 295
      |||

```

```

Db      236 YSNLVGLNDTMENTLAAVDNRNEAETSPETLYAACMENVEPFINGSPONTFVPGILIDA 295
Qy      296 IARNLTIGGDDFKSGOTKKKSVLVDPLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
      |||
      296 IARNLTIGGDDFKSGOTKKKSVLVDPLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
Db      296 IARNLTIGGDDFKSGOTKKKSVLVDPLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
      |||
Qy      356 EISKSNVDDMNNSNAILYEPGEHPDHVVVVIKVPYVGS 395
      |||
      356 EISKSNVDDMNNSNAILYEPGEHPDHVVVVIKVPYVGS 395
Db      356 EISKSNVDDMNNSNAILYEPGEHPDHVVVVIKVPYVGS 395

RESULT 13
ADQ14495
ID      ADQ14495 standard; protein; 510 AA.
XX
AC      ADQ14495;
XX
DT      23-SEP-2004 (first entry)
XX
DE      Mutant soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX
KW      Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW      raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KW      mutant; mutein.
XX
OS      Glycine max.
OS      Synthetic.
XX
XX
Key      Location/Qualifiers
FH      Misc-difference 396
FT      /note= "Wild type Lys substituted by Asn"
XX
XX
XX      US2004128713-A1.
XX
XX      01-JUL-2004.
XX
XX      21-NOV-2003; 2003US-00718952.
XX
XX      08-APR-1997; 97US-00835751.
XX      07-APR-1998; 98WO-US006822.
XX      26-APR-1999; 99US-00299315.
XX      11-MAR-2002; 2002US-00025003.
XX
XX      (HITZ/) HITZ W D.
XX      (SEBA/) SEBASTIAN S A.
XX      (GRAC/) GRACE D J.
XX      (STREI/) STREIT L G.
XX
XX      Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX      WPI; 2004-533135/51.
XX      N-PSDB; ADQ14494.
XX
XX      New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX      useful for producing plants with decreased raffinose, stachyose, and
XX      phytic acid and increased sucrose, leading to valuable and useful soybean
XX      products.
XX
XX      Claim 9; SEQ ID NO 6; 48pp; English.
XX
XX      The invention relates to a nucleic acid fragment encoding a soybean myo-
XX      inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX      synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX      phosphate. The invention also relates to a chimeric gene operably linked
XX      to suitable regulatory sequences, where expression of the chimeric gene
XX      results in a decrease in expression of an endogenous or native gene
XX      encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX      comprising the chimeric gene, a method of making the soybean plant, a
XX      seed of the soybean plant, a soy protein product derived from the
XX      processing of soybean seeds, a method of making or producing a soy
XX      protein product and a method of using a soybean plant homozygous for at
XX      least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX      decreased capacity for the synthesis of myo-inositol 1-phosphate. The

```


CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents a mutant soybean myo-inositol
 CC 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 43.1%; Score 220; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4.7e-211;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPPDFTIAANOEEERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVVLTANTER 235
 DB 176 PLPGIYDPPDFTIAANOEEERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVVLTANTER 235
 QY 236 YSNLVVGLNDTMENTLAAVDREAEISPSSTLYAACWENVPFINGSPONTFVPGILIDA 295
 DB 236 YSNLVVGLNDTMENTLAAVDREAEISPSSTLYAACWENVPFINGSPONTFVPGILIDA 295
 QY 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSX 355
 DB 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSX 355
 QY 356 EISKSNVVDMDVNSNAILYEPGEHPDHVVVIRKVPYVGD 395
 DB 356 EISKSNVVDMDVNSNAILYEPGEHPDHVVVIRKVPYVGD 395

RESULT 14

ADS81998
 ID ADS81998 standard; protein; 510 AA.

AC ADS81998;
 DT 18-NOV-2004 (first entry)

XX Soybean myo-inositol 1-phosphate synthase mutant #1.

DE Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; muten;
 KW phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
 KM flatulence; mutant.

XX Glycine max; line LR33.
 OS Synthetic.

XX Key location/Qualifiers

XX Misc-difference 396 /note= "Wild-type Lys substituted by Arg"

XX US2003074685-A1.

XX 17-APR-2003.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.

XX (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.

XX Hitz WD, Sebastian SA;

XX WPI: 2004-639957/62.

XX N-PsDB; ADS81997.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.

XX Claim 9, SEQ ID NO 6; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences), where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/9, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/9, and a seed sucrose content of
 CC greater than 200 micromol/9, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate), and making or producing a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a mutant myo-
 CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 43.1%; Score 220; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4.7e-211;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPPDFTIAANOEEERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVVLTANTER 235
 DB 176 PLPGIYDPPDFTIAANOEEERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVVLTANTER 235
 QY 236 YSNLVVGLNDTMENTLAAVDREAEISPSSTLYAACWENVPFINGSPONTFVPGILIDA 295
 DB 236 YSNLVVGLNDTMENTLAAVDREAEISPSSTLYAACWENVPFINGSPONTFVPGILIDA 295
 QY 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSX 355
 DB 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSX 355
 QY 356 EISKSNVVDMDVNSNAILYEPGEHPDHVVVIRKVPYVGD 395
 DB 356 EISKSNVVDMDVNSNAILYEPGEHPDHVVVIRKVPYVGD 395

RESULT 15

AAW97882
 ID AAW97882 standard; protein; 510 AA.

XX AAW97882;

XX 07-JUN-1999 (first entry)

XX Maize myo-inositol 1-phosphate synthase.

XX Myo-inositol 1-phosphate synthase; maize; corn; phytate; phytic acid;
 KW transgenic plant; animal nutrition; feedstuff; food.

XX Zea mays.

XX WO9905298-A1.

XX 04-FEB-1999.

PF 17-JUL-1998; 98WO-US014657.
XX
PR 22-JUL-1997; 97US-0053371P.
PR 28-JUL-1997; 97US-0053944P.
PR 08-AUG-1997; 97US-0055526P.
PR 11-AUG-1997; 97US-0055446P.
PR 18-MAY-1998; 98US-0085852P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Martino-Catt SJ, Wang H, Beach LR, Bowen BA, Wang X;
XX
DR WPI: 1999-142948/12.
DR N-PSDB; AAX24407.
XX
XX New polynucleotides controlling phytate metabolism in plants - useful for
PT improving the nutritional content of plants, by enhancing levels of non-
PT phytate phosphorus, and reducing phytate levels.
XX
PS Claim 1a; Page 77-78; 86pp; English.
XX
CC This is the amino acid sequence of maize myo-inositol 1-phosphate
CC synthase, an enzyme involved in the metabolism of phytate. cDNA (see
CC AAX24407) encoding the enzyme was isolated from a maize embryo (15 day
CC post-pollination) cDNA library. Polynucleotides (see AAX24400, AAX24403,
CC AAX24407 and AAX24410-12) encoding maize phosphatidylinositol-3-kinase
CC (see AAW97880), myo-inositol 1,3,4-trisphosphate 5/6-kinase (see
CC AAW97881), myo-inositol 1-phosphate synthase and myo-inositol
CC monophosphatase-3 (see AAW97883), all enzymes involved in phytate
CC metabolism, are claimed. The invention relates to the use of such genes
CC to reduce the levels of phytate, and/or increase the levels of non-
CC phytate phosphorus, in plants used for food or feed. The genes are
CC especially used to improve the nutritional content of plants such as corn
CC and soybean. Transgenic plants, and seed produced by them, are claimed
XX
SQ Sequence 510 AA;

Query Match 14.9%; Score 76; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 7,5e-67;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	369	SNAILYEPGEHPDHVVVVKVVPYVGDSPKAMDEYTSIFPMGKNTIVLHNTCEDSLAAP	428
DB	369	SNAILYEPGEHPDHVVVVKVVPYVGDSPKAMDEYTSIFPMGKNTIVLHNTCEDSLAAP	428
QY	429	IILDVTLAELSTRIQ	444
DB	429	IILDVTLAELSTRIQ	444

Search completed: June 7, 2005, 17:06:33
Job time : 94 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:58:47 ; Search time 24.5 Seconds
(without alignments)
2002.880 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 510
Sequence: 1 MEIENFKVESPWVXKTEETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
1: PIR 79: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	14.9	510	2	T01647 inositol-3-phospha
2	67	13.1	507	2	S52648 inositol-3-phospha
3	67	13.1	509	2	T08436 inositol-3-phospha
4	67	13.1	510	2	T04399 inositol-3-phospha
5	67	13.1	510	2	T50021 inositol-3-phospha
6	67	13.1	510	2	S60302 inositol-3-phospha
7	67	13.1	510	2	D84610 probable myo-inos
8	67	13.1	511	2	T05017 inositol-3-phospha
9	67	13.1	512	2	T12438 inositol-3-phospha
10	48	9.4	511	2	T10964 inositol-3-phospha
11	15	2.9	430	2	T46317 hypothetical prote
12	13	2.5	525	2	T18569 inositol-3-phospha
13	13	2.5	535	2	A30902 inositol-3-phospha
14	12	2.4	520	1	S45452 inositol-3-phospha
15	8	1.6	210	1	B48357 minor outer capsid
16	8	1.6	210	2	S25544 nonstructural prot
17	8	1.6	378	2	F95844 probable dihydroli
18	8	1.6	973	2	AB2340 hypothetical prote
19	7	1.4	42	1	I65746 tropomyosin - huma
20	7	1.4	96	1	WMVZP3 P3 protein - fowlp
21	7	1.4	97	2	JQ0527 capsid assembly pr
22	7	1.4	105	2	B82149 conserved hypothet
23	7	1.4	125	2	AB1995 hypothetical prote
24	7	1.4	147	2	T02273 hypothetical prote
25	7	1.4	157	2	S77352 hypothetical prote
26	7	1.4	178	2	S44910 ZK66.1 protein -
27	7	1.4	191	2	AH0431 probable exported
28	7	1.4	197	2	T15106 hypothetical prote
29	7	1.4	201	2	AH1812 hypothetical prote

30	7	1.4	209	2	A71313 probable V-type AT
31	7	1.4	212	2	T36864 probable cyclohexa
32	7	1.4	231	2	B83111 50S ribosomal prot
33	7	1.4	236	2	T33333 hypothetical prote
34	7	1.4	280	2	T47572 Machado-Joseph dis
35	7	1.4	282	2	AE2618 rRNA methylase (Im
36	7	1.4	290	2	D97400 rRNA methylase (AP
37	7	1.4	292	2	T30321 hypothetical prote
38	7	1.4	294	2	C82497 ribose ABC transpo
39	7	1.4	294	2	AE2457 hypothetical prote
40	7	1.4	299	2	E84262 rhamnosyl transfer
41	7	1.4	300	2	S16430 hypothetical prote
42	7	1.4	308	2	T08796 tropomyosin - huma
43	7	1.4	309	2	C83017 probable short cha
44	7	1.4	309	2	T02613 hypothetical prote
45	7	1.4	311	2	D64209 hydroxymethylgluta

ALIGNMENTS

```

RESULT 1
T01647
inositol-3-phosphate synthase (EC 5.5.1.4) - maize
C:/Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C:/Accession: T01647
R:/Larson, S.R.; Raboy, V.
submitted to the EMBL Data Library, March 1998
A:/Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A:/Reference number: Z14366
A:/Accession: T01647
A:/Status: translated from GB/EMBL/DBJ
A:/Molecule type: mRNA
A:/Residues: 1-510 <IAR>
A:/Cross-references: UNIPROT:Q9FPK7; EMBL:AF056326; NID:g3108052; PIDN:AACT15756.1; PID:g31
A:/Experimental source: strain Early ACR; leaf
C:/Genetics:
A:/Gene: INOI
C:/Function:
A:/Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-ino
A:/Pathway: inositol biosynthesis
A:/Note: NAD cofactor
C:/Keywords: intramolecular lyase; isomerase; NAD
Query Match 14.9%; Score 76; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.7e-70; Mismatches 0; Indels 0; Gaps 0;
Matches 76; Conservative 0;
QY 369 SNAIIYEPGEHDHVVVIKYVPYVDSKRAMDEYTSIFMGKNTIVLHNTCEDSLAAP 428
|||||
DB 369 SNAIIYEPGEHDHVVVIKYVPYVDSKRAMDEYTSIFMGKNTIVLHNTCEDSLAAP 428
QY 429 IILDVLAELSTRIO 444
|||||
DB 429 IILDVLAELSTRIO 444

RESULT 2
S52648
inositol-3-phosphate synthase (EC 5.5.1.4) - Citrus paradisi
C:/Species: Citrus paradisi
C:/Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:/Accession: S52648
R:/Holland, D.
submitted to the EMBL Data Library, April 1994
A:/Reference number: S52648
A:/Accession: S52648
A:/Molecule type: DNA
A:/Residues: 1-507 <HOL>
A:/Cross-references: UNIPROT:P42802; GB:Z32632; NID:g602564; PIDN:CAA83565.1; PID:g602565
C:/Genetics:

```

A:Gene: INO1
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase, isomerase

Query Match 13.1%; Score 67; DB 2; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.5e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360

QY 361 NVVDDMV 367
DB 361 NVVDDMV 367

RESULT 3
T08436
inositol-3-phosphate synthase (EC 5.5.1.4) [similarity] - rape
C:Species: Brassica napus (rape)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08436
R:Hussein, A.; Bourgeois, J.; Polvi, S.; Tsang, E.; Keller, W.A.; Georges, F.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z16418
A:Accession: T08436
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-509 <HUS>
A:Cross-references: UNIPROT:Q96348; EMBL:U66307; NID:G1513227; PID:G1513228
C:Function:
A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase, isomerase

Query Match 13.1%; Score 67; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.5e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTRSKSEIKS 359

QY 361 NVVDDMV 367
DB 361 NVVDDMV 367

RESULT 4
T04399
inositol-3-phosphate synthase (EC 5.5.1.4) - barley
C:Species: Hordeum vulgare (barley)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04399
R:Larson, S.R.; Raboy, V.
submitted to the EMBL Data Library, March 1998
A:Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A:Reference number: Z13366
A:Accession: T04399
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-510 <LAR>
A:Cross-references: UNIPROT:O65195; EMBL:AF056325; NID:G3152730; PID:AA017133.1; PID:G3
A:Experimental source: cv. Harrington
C:Genetics:
A:Gene: INO1
A:Map position: 4
C:Function:
A:Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inc
A:Pathway: inositol biosynthesis
A>Note: first step
C:Superfamily: myo-inositol-1-phosphate synthase

C:Keywords: intramolecular lyase, isomerase, NAD

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.5e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360

QY 361 NVVDDMV 367
DB 361 NVVDDMV 367

RESULT 5
T50021
inositol-3-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis thaliana
N:Alternate names: protein T31P16.160
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50021
R:Bevan, M.; Zimmermann, W.; Grenetisen, A.; Wandut, R.; Kalicki, J.; Woldmann, P.; Smi
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25027
A:Accession: T50021
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <BEV>
A:Cross-references: UNIPROT:O9LX12; EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.160
A:Experimental source: cultivar Columbia; BAC clone T31P16
C:Genetics:
A:Gene: ATSP:T31P16.160
A:Map position: 5
A:Introns: 63/2; 86/2; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase, isomerase, NAD

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.5e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360

QY 361 NVVDDMV 367
DB 361 NVVDDMV 367

RESULT 6
S60302
inositol-3-phosphate synthase (EC 5.5.1.4) - Spirodela polyrrhiza
C:Species: Spirodela polyrrhiza
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60302
R:Smart, C.C.; Fleming, A.J.
Plant J. 4, 279-293, 1993
A:Title: A plant gene with homology to D-myo-inositol-3-phosphate synthase is rapidly an
A:Reference number: S60302; MUID:94035182; PMID:8220483
A:Accession: S60302
A:Molecule type: mRNA
A:Residues: 1-510 <SMA>
A:Cross-references: UNIPROT:P42803; EMBL:Z11693; NID:G396209; PID:CAA77751.1; PID:G55864
A:Genetics:
A:Gene: turl
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase, isomerase

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.5e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 301 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367
 RESULT 7
 D84610
 Probable myo-inositol 1-phosphate synthase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004
 C:Accession: D84610
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beattie, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: D84610
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-510 <STO>
 A:Cross-references: UNIPROT:Q38862; GB:AE002093; NID:g4567202; PIDN:AAD23618.1; GSPDB:GN
 A:Gene: At2g22240
 A:Map position: 2
 C:Superfamily: Myo-inositol-1-phosphate synthase
 Query Match 13.1%; Score 67; DB 2; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 301 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367
 RESULT 8
 T05017
 Inositol-3-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thaliana
 N:Alternate names: protein T19P19.190
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C:Accession: T05017
 R:Bevan, M.; Monfort, A.; Casauberta, E.; Puigdomenech, P.; Hohseisel, J.; Mewes, H.W.;
 submitted to the Protein Sequence Database, April 1998
 A:Reference number: Z15394
 A:Accession: T05017
 A:Molecule type: DNA
 A:Residues: 1-511 <BEV>
 A:Cross-references: UNIPROT:P42801; EMBL:AL022605
 A:Experimental source: cultivar Columbia; BAC clone T19P19
 C:Function:
 A:Map position: 4
 A:Introns: 64/2; 87/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3
 A>Note: T19P19.190
 C:Superfamily: myo-inositol-1-phosphate synthase
 C:Keywords: intramolecular lyase; isomerase; NAD
 Query Match 13.1%; Score 67; DB 2; Length 511;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 301 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360

DB 302 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 361
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367
 RESULT 9
 T12438
 Inositol-3-phosphate synthase (EC 5.5.1.4) - common ice plant
 C:Species: Mesembryanthemum crystallinum (common ice plant)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T12438
 R:Rishitani, M.; Majumder, A.L.; Bornhauser, A.; Michalowski, C.B.; Jensen, R.G.; Bohmert,
 Plant J. 9, 537-548, 1996
 A>Title: Coordinate transcriptional induction of myo-inositol metabolism during environme
 A:Reference number: Z17518; MUID:96208959; PMID:8624516
 A:Accession: T12438
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-512 <ISH>
 A:Cross-references: UNIPROT:Q40271; EMBL:U32511; NID:g975887; PIDN:AAB03687.1; PID:g97588
 C:Superfamily: myo-inositol-1-phosphate synthase
 C:Keywords: intramolecular lyase; isomerase
 Query Match 13.1%; Score 67; DB 2; Length 512;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 301 LIGGDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKS 362
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367
 RESULT 10
 T10964
 Inositol-3-phosphate synthase (EC 5.5.1.4) - kidney bean
 N:Alternate names: IL-myo-inositol 1-phosphate synthase
 C:Species: Phaseolus vulgaris (kidney bean)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T10964
 R:Wang, X.; Johnson, M.D.
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z17234
 A:Accession: T10964
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-511 <MAN>
 A:Cross-references: UNIPROT:Q41107; EMBL:U38920; NID:g1066282; PID:g1066283
 A:Experimental source: strain Taylor's horticultural; root
 C:Function:
 A:Description: catalyzes reversible conversion of D-glucose 6-phosphate to IL-myo-inosit
 A:Pathway: myo-inositol biosynthesis
 C:Superfamily: myo-inositol-1-phosphate synthase
 C:Keywords: intramolecular lyase; isomerase; NAD
 Query Match 9.4%; Score 48; DB 2; Length 511;
 Best Local Similarity 100.0%; Pred. No. 6.4e-41;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 320 DFLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKSNVVDDMV 367
 DB 321 DFLVAGIKRPTISVSYNHLGNNDGMNLSAPQTRSKSEIKSNVVDDMV 368
 RESULT 11
 T46317
 Hypothetical protein DKFZp434A0612.1 - human
 C:Species: Homo sapiens (man)

A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo
A:Note: requires NAD
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 2.4%; Score 12; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 INGSPOINTFPG 290
Db 285 INGSPOINTFPG 296

RESULT 15

B48357
minor outer capsid protein - porcine rotavirus C (strain Cowden)
N:Alternate names: nonstructural protein NS26
C:Species: porcine rotavirus C
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: B48357
R:Bremont, M.; Chabanne-Vaucherot, D.; Cohen, J.
Arch. Virol. 130, 85-92, 1993
A:Title: Sequence analysis of three non structural proteins of a porcine group C (Cowden
A:Reference number: A48357; MUID:93277387; PMID:8389118
A:Accession: B48357
A:Molecule type: genomic RNA
A:Residues: 1-210 <BRE>
A:Cross-references: UNIPROT:P36358
C:Genetics:
A:Map position: segment 10
C:Superfamily: rotavirus minor outer capsid protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:30,120/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.6%; Score 8; DB 1; Length 210;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 313 KMKSVLVD 320
Db 198 KMKSVLVD 205

Search completed: June 7, 2005, 17:10:36
Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:49:04 ; Search time 86 Seconds
(without alignments)
3036.749 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 510

Sequence: 1 MEIENPKVSPVVKYETETI.....NIMBACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Listing first 45 summaries

Database : UniProt 03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	236	46.3	510	2	094C02 glycine max
2	236	46.3	510	2	094K03 glycine max
3	32	18.0	472	2	094G22 phaseolus v
4	76	14.9	510	1	INO1_MAIZE
5	76	14.9	510	1	INO1_NICPA
6	76	14.9	510	1	INO1_TOBAC
7	76	14.9	512	2	07XZ66 nicotiana t
8	73	14.3	510	2	09AV99 porteresia
9	72	14.1	510	2	094G38 avena sativ
10	67	13.1	364	2	09AWG8 suada sala
11	67	13.1	443	2	09AWG8 actinidia a
12	67	13.1	443	2	08H1B7 lolium pere
13	67	13.1	509	2	INO1_CITPA
14	67	13.1	509	2	09AR12 avicennia m
15	67	13.1	510	1	INO1_BRANA
16	67	13.1	510	1	INO1_HORVU
17	67	13.1	510	1	INO1_ORYSA
18	67	13.1	510	1	INO1_ORYSA
19	67	13.1	510	1	INO1_SSSIN
20	67	13.1	510	1	INO1_SPIPO
21	67	13.1	510	1	INO1_ARATH
22	67	13.1	510	1	INO1_ARATH
23	67	13.1	511	1	INO1_ARATH
24	67	13.1	512	1	INO1_ARATH
25	55	10.8	509	2	08SSN2 mesembryant
26	48	9.4	511	1	INO1_PRAVU
27	47	9.2	164	2	09FT40 phaseolus v
28	38	7.5	186	2	042548 lycopersico
29	36	7.1	409	2	084SA1 arabidopsis
30	29	5.7	565	1	INO1_DROME
31	27	5.3	135	2	09AT03 solanum tub

32	27	5.3	166	2	09FUP2 lycopersico
33	23	4.5	150	2	095U04 branchiosteo
34	23	4.5	560	2	06DDT1 xenopus lae
35	23	4.5	563	2	07ZX10 xenopus lae
36	21	4.1	561	2	07P2B9 anopheles g
37	17	3.3	526	2	07Z525 homo sapien
38	17	3.3	557	2	09JHU9 myo-inosi
39	17	3.3	558	2	06NXT5 homo sapien
40	17	3.3	558	2	09H2Y2 homo sapien
41	17	3.3	558	2	09NPH2 homo sapien
42	17	3.3	558	2	09NPH2 homo sapien
43	16	3.1	220	2	06UP00 aspergillus
44	16	3.1	537	2	087IUS neurospora
45	15	2.9	181	2	09BR65 homo sapien

ALIGNMENTS

RESULT 1	ID	094C02	PRELIMINARY;	PRT;	510 AA.
AC	094C02;	01-DEC-2001 (TREMBlrel. 19, Created)			
DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)				
DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)				
DE	Myo-inositol-1-phosphate synthase (EC 5.5.1.4).				
OS	Glycine max (Soybean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
OX	NCBI_TaxId=3847;				
RN	(1)				
RA	SEQUENCE FROM N.A.				
RP	Hitc W.D., Carlson T.J., Kerr P., Sebastian S.;				
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY038802; AKK72098.1; -				
DR	HSSP; P12986; 1P1H.				
DR	GO; GO:004512; F:inositol-3-phosphate synthase activity; IEA.				
DR	GO; GO:0016853; F:isomerase activity; IEA.				
DR	GO; GO:0006021; P:myo-inositol biosynthesis; IEA.				
DR	GO; GO:000654; P:phospholipid biosynthesis; IEA.				
DR	InterPro: IPR002587; Inos-1-P_synth.				
DR	Pfam: PF01658; Inos-1-P_synth; 1.				
KW	Isomerase.				
SQ	SEQUENCE 510 AA; 56475 MW; A1E4C7F1643918E CRC64;				
Query Match	46.3%; Score 236; DB 2; Length 510;				
Best Local Similarity	100.0%; Pred. No. 4.8e-238; Indels 0; Gaps 0;				
Matches	236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	176 PLPGIYDDPFIANOEERANNVIGTKQBOVOQIIKIDIKAFKEATKVDKVVVLMWTANTER	235			
DB	176 PLPGIYDDPFIANOEERANNVIGTKQBOVOQIIKIDIKAFKEATKVDKVVVLMWTANTER	235			
QY	236 YSNLVGGLNDTMENTLLAANDRNEAISTPSTLYAIAICWENVPFINGSPONTVPGLIDDA	295			
DB	236 YSNLVGGLNDTMENTLLAANDRNEAISTPSTLYAIAICWENVPFINGSPONTVPGLIDDA	295			
QY	296 IARNTLIGSDPDKSGOTKAKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNTLSAPOTPRSK	355			
DB	296 IARNTLIGSDPDKSGOTKAKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNTLSAPOTPRSK	355			
QY	356 EISKSNVDDMVNSNAIYEPGEHPDHVVIKVVPYVGSKRAMDYTSSEIFMGK 411				
DB	356 EISKSNVDDMVNSNAIYEPGEHPDHVVIKVVPYVGSKRAMDYTSSEIFMGK 411				
RESULT 2					
ID	094K03	PRELIMINARY;	PRT;	510 AA.	
AC	094K03;	01-DEC-2001 (TREMBlrel. 19, Created)			

```
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Myo-inositol-3-phosphate synthase.
CN Name=MIPS;
OS Glycine max (soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21196082; PubMed=11299373; DOI=10.1104/pp.125.4.1941;
RA Hegeman C.E., Good L.L., Grabau E.A.;
RT "Expression of D-myo-inositol-3-phosphate synthase in soybean.
RT Implications for phytic acid biosynthesis."
RL Plant Physiol. 125:1941-1948(2001).
DR EMBL; AF293970; AAK4986.1; -.
DR HSSP; P11986; 1PIH.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
DR NON_TER 472 472
SQ SEQUENCE 472 AA; 52296 MW; 237204E1A370560F CRC64;
```

```
Query Match 18.0%; Score 92; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.9e-87;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGDDDFSGGQTKKSVLVDFLVAGIKPTSTIVYNHGNNDGNLSPOTFRSKETS 360
DB 301 LIGDDDFSGGQTKKSVLVDFLVAGIKPTSTIVYNHGNNDGNLSPOTFRSKETS 360
QY 361 NVVDDMVNSNAILYEPGEHPDHVVVVIKYPYV 392
DB 361 NVVDDMVNSNAILYEPGEHPDHVVVVIKYPYV 392

RESULT 4
INOI_MAIZE
ID INOI_MAIZE STANDARD; PRT; 510 AA.
AC Q9PRK7; OCS196;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (MI-1-P synthase) (IPS).
OS Zea mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Early ACR; TISSUE=leaf;
RA Larson S.R., Raboy V.;
RT "Linkage mapping maize and barley myo-inositol 1-phosphate synthase
RT genes."
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shukla S., Vartosi T.T.;
RT "Genomic sequence of maize myo-inositol 1-phosphate synthase gene."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: MAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF056326; AAC15756.1; -.
DR EMBL; AF323175; AAG40328.1; -.
DR PIR; T01647; T01647.
DR HSSP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
KW CONFLICT 18 M -> T (in Ref. 2).
FT CONFLICT 351 A -> T (in Ref. 2).
SQ SEQUENCE 510 AA; 56245 MW; DA59EFC391CB6D CRC64;

Query Match 14.9%; Score 76; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.1e-70;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAILYEPGEHPDHVVVVIKYPYVGDSCRAMDEYSEIFMGKNTIVLHNTCESSLAP 428
DB 369 SNAILYEPGEHPDHVVVVIKYPYVGDSCRAMDEYSEIFMGKNTIVLHNTCESSLAP 428
```

QY 429 ILLDLVLAELSTRIO 444
 DB 429 ILLDLVLAELSTRIO 444

RESULT 5
 INOI_NICPA STANDARD; PRT; 510 AA.

AC 09SSV4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (Mt-1-P synthase) (IPS).
 GN Name=INPS1;
 OS Nicotiana paniculata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Nicotiana.
 CC NCBI_TaxId=62141;
 RX SEQUENCE FROM N.A.
 RA Hashimoto A., Yamada S., Komori T.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; AB032073; BAA84084.1; -
 DR HSSP; P11986; 1PIH.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56385 MW; 415B81C27A267666 CRC64;

Query Match 14.9%; Score 76; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 3.1e-70;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIIYEGEHDPHVVIKYVPYVGDSCRAMDEYSEIFMGKNTIVLHNTCEDSLAAP 428
 DB 369 SNAIIYEGEHDPHVVIKYVPYVGDSCRAMDEYSEIFMGKNTIVLHNTCEDSLAAP 428

QY 429 ILLDLVLAELSTRIO 444
 DB 429 ILLDLVLAELSTRIO 444

RESULT 6
 INOI_TOBAC STANDARD; PRT; 510 AA.

AC 09LW96;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (Mt-1-P synthase) (IPS).
 GN Nicotiana tabacum (Common tobacco).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Solanales; Solanaceae; Nicotiana.

OC Lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxId=4097;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=20399434; PubMed=10945337;
 RA Hara K., Yagi M., Koizumi N., Kusano T., Sano H.;
 RT "Screening of wound-responsive genes identifies an immediate-early expressed gene encoding a highly charged protein in mechanically wounded tobacco plants."
 RL Plant Cell Physiol. 41:684-691(2000).
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; AB009881; BAA95788.1; -
 DR HSSP; P11986; 1PIH.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56369 MW; 4BA8FDDA5BFE6D4D CRC64;

Query Match 14.9%; Score 76; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 3.1e-70;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIIYEGEHDPHVVIKYVPYVGDSCRAMDEYSEIFMGKNTIVLHNTCEDSLAAP 428
 DB 369 SNAIIYEGEHDPHVVIKYVPYVGDSCRAMDEYSEIFMGKNTIVLHNTCEDSLAAP 428

QY 429 ILLDLVLAELSTRIO 444
 DB 429 ILLDLVLAELSTRIO 444

RESULT 7
 QYXZ6 PRELIMINARY; PRT; 512 AA.

AC 07XZ6;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Inositol 1-phosphate synthase (EC 5.5.1.4).
 GN Name=PINOL;
 OS Porteresia coarctata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrharioideae; Oryzaceae; Porteresia.
 CC NCBI_TaxId=77588;
 RX SEQUENCE FROM N.A.
 RA MEDLINE=22912548; PubMed=14550537; DOI=10.1016/S0014-5793(03)00974-8;
 RA Majumder A.L., Chatterjee A., Ghosh Dasidhar K., Majee M.;
 RT "Diversification and evolution of L-myo-inositol 1-phosphate synthase."
 RL FEBS Lett. 553:3-10(2003).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Majumder A.N.L., Majee M.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

```
RP SEQUENCE FROM N.A.
RA Majumder A.N.L., Majee M.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF412340; AAF74579.1; -.
DR HSSP; P11986; 1PKF.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1_P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
KW Isomerase.
SQ SEQUENCE 512 AA; 56651 MW; 69402D547EE1D7F6 CRC64;

Query Match 14.9%; Score 76; DB 2; Length 512;
Best Local Similarity 100.0%; Pred. No. 3.2e-70;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIYEPGEHPDHVVVVKYVPYVGDGSKRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAPTLL 428
DB 371 SNAIYEPGEHPDHVVVVKYVPYVGDGSKRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAPTLL 430
QY 429 IILDVLAELSTRIQ 444
DB 431 IILDVLAELSTRIQ 446

RESULT 8
Q9AV99 PRELIMINARY; PRT; 510 AA.
ID Q9AV99;
AC Q9AV99;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Myo-inositol-1-phosphate synthase.
GN Name=MTP;
OS Avena sativa (Oat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Avenae; Avena.
OC NCBI_TaxID=4498;
OK [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Seed;
RC MEDLINE=99097040; PubMed=9880347; DOI=10.1104/pp.119.1.65;
RX Yoshida Y.T., Wada T., Koyama H., Mizobuchi R., Naito S.;
RT "Temporal and spatial patterns of accumulation of the transcript of
RT Myo-inositol-1-phosphate synthase and phytin-containing particles
RT during seed development in rice.";
RL Plant Physiol. 119:65-72(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RA Toyonaga D., Sanecka H.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB059557; BAB40956.2; -.
DR HSSP; P11986; 1PK.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 510 AA; 56130 MW; A0F9D9F80CA0C6CC CRC64;

Query Match 14.3%; Score 73; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.4e-67;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 372 ILYEPGEHPDHVVVVKYVPYVGDGSKRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAPTLL 431
DB 372 ILYEPGEHPDHVVVVKYVPYVGDGSKRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAPTLL 431
QY 432 DLVLAELSTRIQ 444
DB 432 DLVLAELSTRIQ 444
```

```
DB 432 DLVLAELSTRIQ 444

RESULT 9
Q944C3 PRELIMINARY; PRT; 510 AA.
ID Q944C3;
AC Q944C3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Myo-inositol-1-phosphate synthase.
GN Name=INPS;
OS Suaeda salsa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Suaeda.
OK NCBI_TaxID=126914;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang J.P., Wang P.P., Sun Y.F., Zhao Y.X., Zhang H.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF433879; AAL28131.1; -.
DR HSSP; P11986; 1PK.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 510 AA; 56665 MW; 6C31006D2E1C508A CRC64;

Query Match 14.1%; Score 72; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.9e-66;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 372 ILYEPGEHPDHVVVVKYVPYVGDGSKRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAPTLL 431
DB 373 ILYEPGEHPDHVVVVKYVPYVGDGSKRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAPTLL 432
QY 432 DLVLAELSTRIQ 443
DB 432 DLVLAELSTRIQ 443
QY 433 DLVLAELSTRIQ 444
DB 433 DLVLAELSTRIQ 444

RESULT 10
Q9AWG8 PRELIMINARY; PRT; 364 AA.
ID Q9AWG8;
AC Q9AWG8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Myo-inositol-1-phosphate synthase.
OS Actinidia arguta.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Ericales; Actinidiaceae; Actinidia.
OK NCBI_TaxID=64478;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Young mature leaf;
RA Klages K., Fitzgerald A., Moodie M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY005128; AAF97409.1; -.
DR HSSP; P11986; 1PK.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 364 AA; 40246 MW; 1CB64D6FF78127D CRC64;

Query Match 13.1%; Score 67; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 6.4e-61;
```

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSISKS 360
 DB 155 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSISKS 214

QY 361 NVVDDMV 367
 DB 215 NVVDDMV 221

RESULT 11
 08H1B7 PRELIMINARY; PRT; 443 AA.

AC 08H1B7
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Myo-inositol phosphate synthase (Fragment).
 OS Lolium perenne (Perennial ryegrass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Poaceae; Lolium.
 OC NCB1_TaxId=4522;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Amiard V., Prud'homme M.-P., Le Dantec C.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY154382; JAKN52772.1; --
 DR HSSP; P11986; 1PKF.
 DR GO; GO:0004512; P:myo-inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR NON TER
 FT 443 443
 SQ SEQUENCE 443 AA; 48572 MW; E8BB1A03779EA9FA CRC64;

Query Match 13.1%; Score 67; DB 2; Length 443;
 Best Local Similarity 100.0%; Pred. No. 7.6e-61;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSISKS 360
 DB 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSISKS 360

QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 12
 ID INOI CITPA STANDARD; PRT; 507 AA.

AC P42802;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
 synthase) (MI-1-P synthase) (IPS).
 OS Citrus paradisi (Grapefruit).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids 1; Sapindales; Rutaceae; Citrus.
 OC NCB1_TaxId=37656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-leaf;
 RC MEDLINE=95148748; PubMed=7846170; DOI=10.1104/pp.106.4.1689;
 RA Abu-Abied M., Holland D.;
 RT "The gene c-inol from Citrus paradisi is highly homologous to curt1
 RT inoi from Yeast and Spirodela encoding for myo-inositol phosphate
 RT synthase.";

RL Plant Physiol. 106:1689-1689 (1994).
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
 CC phosphate.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z32632; CAAB3565.1; --
 DR PIR; S52648; S52648.
 DR HSSP; P11986; 1PKF.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 SQ SEQUENCE 507 AA; 56334 MW; 45D7892891BDFD8 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 507;
 Best Local Similarity 100.0%; Pred. No. 8.6e-61;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSISKS 360
 DB 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSISKS 360

QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 13
 ID 09AR12 PRELIMINARY; PRT; 509 AA.

AC 09AR12;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Myo-inositol 1-phosphate synthase.
 OS Avicennia marina (Grey mangrove).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Lamiales; Acanthaceae; Acanthaceae incertae sedis; Avicennia.
 OC NCB1_TaxId=82927;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jithesh M.N., Parani M., Parida A.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY028259; AAK21969.1; --
 DR HSSP; P11986; 1PKF.
 DR GO; GO:0004512; P:myo-inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 SQ SEQUENCE 509 AA; 55978 MW; 23C8D354BAF3BD0F CRC64;

Query Match 13.1%; Score 67; DB 2; Length 509;
 Best Local Similarity 100.0%; Pred. No. 8.6e-61;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSISKS 360
 DB 300 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSISKS 359

QY 361 NVVDDMV 367

Db 360 NVVDDMV 366

RESULT 14

INOL_BRANA STANDARD; PRT; 510 AA.
 ID INOL_BRANA
 AC 096348;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 OS Brassica napus (rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 NCBI_TaxID=3708;
 OK NCB1
 RN
 RP SEQUENCE FROM N.A.
 RA Hussain A., Bourgeois J., Polvi S., Tsang E., Keller W.A., Georges F.;
 RT "Cloning of a full length cDNA encoding myo-inositol 1-phosphate synthase from Brassica napus." to the EMBL/GenBank/DBJ databases.
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U66307; AAB06756.2; -.
 DR PIR; T08436; T08436.
 DR HSSP; P11986; P11986.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW SEQUENCE 510 AA; 56377 MW; A40EB6558D80739 CRC64;
 SQ
 Query Match 13.1%; Score 67; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 8.7e-61;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIGGDDFKSGQTKMKSIVLDFVGAGIKPTSIIVSYNHLGNNDGNMLSAPOFRSKSISKS 360
 DB 301 LIGGDDFKSGQTKMKSIVLDFVGAGIKPTSIIVSYNHLGNNDGNMLSAPOFRSKSISKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367
 RESULT 15
 INOL_HORVU STANDARD; PRT; 510 AA.
 ID INOL_HORVU
 AC 065195;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
 NCBI_TaxID=4513;
 OK NCB1
 RN
 RP SEQUENCE FROM N.A.
 RA Larson S.R., Raboy V.;
 RT "Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF056325; AAC17133.1; -.
 DR PIR; T04339; T04339.
 DR HSSP; P11986; P11986.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW SEQUENCE 510 AA; 56173 MW; EA63138121692724 CRC64;
 SQ
 Query Match 13.1%; Score 67; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 8.7e-61;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIGGDDFKSGQTKMKSIVLDFVGAGIKPTSIIVSYNHLGNNDGNMLSAPOFRSKSISKS 360
 DB 301 LIGGDDFKSGQTKMKSIVLDFVGAGIKPTSIIVSYNHLGNNDGNMLSAPOFRSKSISKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367
 Search completed: June 7, 2005, 17:09:34
 Tool time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:00:20 ; Search time 31 Seconds
(without alignments) 1228.098 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 510
Sequence: 1 MFIEFKYBSPVKYTEREI.....NIMRACVGLAPENNILRYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	14.9	510	3	US-09-118-442-11 Sequence 11, Appl
2	76	14.9	510	3	US-09-677-064-11 Sequence 11, Appl
3	67	13.1	510	4	US-09-727-628-2 Sequence 2, Appl
4	13	2.5	533	4	US-09-734-237B-73 Sequence 73, Appl
5	13	2.5	534	4	US-09-734-237B-75 Sequence 75, Appl
6	12	2.4	525	4	US-09-248-796A-17234 Sequence 17234, A
7	1481	9	1481	4	US-09-231-899-70 Sequence 70, Appl
8	7	1.4	55	4	US-09-640-211A-2116 Sequence 2116, Ap
9	7	1.4	93	4	US-09-902-540-15884 Sequence 15884, A
10	7	1.4	134	4	US-09-270-767-34609 Sequence 34609, A
11	7	1.4	134	4	US-09-270-767-49826 Sequence 49826, A
12	7	1.4	160	4	US-09-957-641A-19 Sequence 19, Appl
13	7	1.4	191	4	US-09-252-991A-24024 Sequence 24024, A
14	7	1.4	224	2	US-08-272-255-16 Sequence 16, Appl
15	7	1.4	224	5	PCT-US95-08565-16 Sequence 16, Appl
16	7	1.4	254	4	US-09-252-991A-21339 Sequence 21339, A
17	7	1.4	269	4	US-09-902-540-16003 Sequence 16003, A
18	7	1.4	273	3	US-08-235-836C-142 Sequence 142, App
19	7	1.4	273	3	US-08-235-836C-144 Sequence 144, App
20	7	1.4	284	4	US-09-914-259-62 Sequence 62, Appl
21	7	1.4	284	4	US-10-164-595-32 Sequence 32, Appl
22	7	1.4	295	4	US-09-270-767-57050 Sequence 57050, A
23	7	1.4	295	4	US-09-270-767-57401 Sequence 57401, A
24	7	1.4	306	4	US-09-107-532A-5995 Sequence 5995, Ap
25	7	1.4	308	4	US-09-489-039A-11598 Sequence 11598, A
26	7	1.4	322	4	US-09-565-501A-112 Sequence 112, App
27	7	1.4	322	4	US-09-639-206A-112 Sequence 112, App

28	7	1.4	322	4	US-09-874-923-112	Sequence 112, App
29	7	1.4	340	4	US-09-489-039A-9804	Sequence 9804, Ap
30	7	1.4	361	4	US-09-489-039A-9936	Sequence 9936, Ap
31	7	1.4	361	4	US-09-710-279-160	Sequence 160, App
32	7	1.4	364	4	US-09-902-540-16435	Sequence 16435, A
33	7	1.4	381	4	US-09-254-776B-53	Sequence 53, Appl
34	7	1.4	388	4	US-09-949-016-7631	Sequence 7631, Ap
35	7	1.4	399	4	US-09-543-681A-5325	Sequence 5325, Ap
36	7	1.4	443	4	US-09-270-767-46616	Sequence 46616, A
37	7	1.4	443	4	US-09-949-016-11613	Sequence 11613, A
38	7	1.4	446	4	US-09-252-991A-22844	Sequence 22844, A
39	7	1.4	467	4	US-09-489-039A-12726	Sequence 12726, A
40	7	1.4	472	4	US-09-489-039A-8121	Sequence 8121, Ap
41	7	1.4	472	4	US-09-134-000C-3893	Sequence 3893, Ap
42	7	1.4	548	4	US-09-902-540-11870	Sequence 11870, A
43	7	1.4	557	3	US-08-979-608A-5	Sequence 5, Appl
44	7	1.4	557	4	US-09-517-849-5	Sequence 5, Appl
45	7	1.4	557	4	US-09-616-289-5	Sequence 5, Appl

ALIGNMENTS

```
RESULT 1
US-09-118-442-11
; Sequence 11, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-118-442-11
;
;
Query Match          14.9% Score 76; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 4,4e-79;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 369 SNAIYBGEHPDHVVIKYPYVGDSTRANDDEYTSKIFMGCKTIVLANTCEDSLAP 428
DB 369 SNAIYBGEHPDHVVIKYPYVGDSTRANDDEYTSKIFMGCKTIVLANTCEDSLAP 428
QY 429 IILDVLLAESTRIO 444
DB 429 IILDVLLAESTRIO 444
RESULT 2
US-09-677-064-11
; Sequence 11, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
```

```
/ TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
/ TITLE OF INVENTION: Plants and Uses Thereof
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677,064
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ PRIOR FILING DATE: 1998-07-17
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Zea mays
US-09-677-064-11
```

```
Query Match 14.9%; Score 76; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.4e-70;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 369 SNAIYEPGEHPDHVVIKYVPYVGDSSKRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAP 428
DB 369 SNAIYEPGEHPDHVVIKYVPYVGDSSKRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAP 428
QY 429 IILDVLVLAELSTRIQ 444
DB 429 IILDVLVLAELSTRIQ 444
```

```
RESULT 3
US-09-727-628-2
/ Sequence 2, Application US/09727628
/ Patent No. 6791013
/ GENERAL INFORMATION:
/ APPLICANT: Armstrong, Katherine
/ APPLICANT: Hey, Timothy D
/ APPLICANT: Folkerts, Otto
/ APPLICANT: Smith, Kelley A
/ APPLICANT: Hopkins, Nicole L
/ TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
/ FILE REFERENCE: 50597
/ CURRENT APPLICATION NUMBER: US/09/727,628
/ CURRENT FILING DATE: 2000-12-01
/ PRIOR APPLICATION NUMBER: US 60/168,612
/ PRIOR FILING DATE: 1999-12-02
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Zea mays
US-09-727-628-2
```

```
Query Match 13.1%; Score 67; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 9.7e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 301 LIGGDFRSQGTQKMSVLVDFVGAGIKPTSIVSYNHLGNNNGMNLAPQFRSKSISKS 360
DB 301 LIGGDFRSQGTQKMSVLVDFVGAGIKPTSIVSYNHLGNNNGMNLAPQFRSKSISKS 360
QY 361 NVVDDMV 367
DB 361 NVVDDMV 367
```

```
RESULT 4
US-09-734-237B-73
```

```
/ Sequence 73, Application US/09734237B
/ Patent No. 6818752
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bul, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ CURRENT FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 73
/ LENGTH: 533
/ TYPE: PRT
/ ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-73
```

```
Query Match 2.5%; Score 13; DB 4; Length 533;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 279 INGSPONTFVPGL 291
DB 293 INGSPONTFVPGL 305
```

```
RESULT 5
US-09-734-237B-75
/ Sequence 75, Application US/09734237B
/ Patent No. 6818752
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bul, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ CURRENT FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 75
/ LENGTH: 534
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic protein derived from Saccharomyces cerevisiae myo-inositol-1-phosphate synthase, having a glycine residue inserted after
/ OTHER INFORMATION: the initiating methionine
US-09-734-237B-75
```

```
Query Match 2.5%; Score 13; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 279 INGSPONTFVPGL 291
DB 294 INGSPONTFVPGL 306
```

```
RESULT 6
US-09-248-796A-17234
/ Sequence 17234, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
/ FILE REFERENCE: 107196.132
```

```

; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17234
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17234

```

```

Query Match      2.4%; Score 12; DB 4; Length 525;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      279 INGSFONTEVPG 290
        |||||
Db      290 INGSFONTEVPG 301

```

```

RESULT 7
US-09-231-899-70
; Sequence 70, Application US/09231899
; Patent No. 6566583
; GENERAL INFORMATION:
; APPLICANT: Lasner, Michael
; APPLICANT: Metz, James G
; APPLICANT: Facclotti, Daniel
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKS GENES
; FILE REFERENCE: CGNE.131.02US
; CURRENT APPLICATION NUMBER: US/09/231,899
; EARLIER FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: 09/030,793
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 70
; LENGTH: 1481
; TYPE: PRT
; ORGANISM: Schizochytrium aggregatum
US-09-231-899-70

```

```

Query Match      1.8%; Score 9; DB 4; Length 1481;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      248 ENLLAAVDR 256
        |||||
Db      1092 ENLLAAVDR 1100

```

```

RESULT 8
US-09-640-211A-2116
; Sequence 2116, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Modifications and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2116
; LENGTH: 55

```

```

; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-2116

```

```

Query Match      1.4%; Score 7; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      485 ORAMLEN 491
        |||||
Db      36 ORAMLEN 42

```

```

RESULT 9
US-09-902-540-15484
; Sequence 15484, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15484
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15484

```

```

Query Match      1.4%; Score 7; DB 4; Length 93;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      473 PPPTPVV 479
        |||||
Db      54 PPPTPVV 60

```

```

RESULT 10
US-09-270-767-34609
; Sequence 34609, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34609
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34609

```

```

Query Match      1.4%; Score 7; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      286 TFVPGLI 292
        |||||
Db      99 TFVPGLI 105

```

```
RESULT 11
US-09-270-767-49826
; Sequence 49826, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49826
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49826

Query Match
Best Local Similarity 1.4%; Score 7; DB 4; Length 134;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 TPVPGI 292
Db 99 TPVPGI 105

RESULT 12
US-09-957-641A-19
; Sequence 19, Application US/09957641A
; Patent No. 6770744
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-00 US
; CURRENT APPLICATION NUMBER: US/09/957,641A
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/234,047
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/236,460
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Porcine
US-09-957-641A-19

Query Match
Best Local Similarity 1.4%; Score 7; DB 4; Length 160;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 476 TPVNAL 482
Db 119 TPVNAL 125

RESULT 13
US-09-252-991A-24024
; Sequence 24024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.116
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24024
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24024

Query Match
Best Local Similarity 1.4%; Score 7; DB 4; Length 191;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 APVPPG 475
Db 85 APVPPG 91

RESULT 14
US-08-272-255-16
; Sequence 16, Application US/08272255
; Patent No. 5824859
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859r18
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/272,255
; FILING DATE: 08-JUL-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: UPN-1795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-272-255-16

Query Match
Best Local Similarity 1.4%; Score 7; DB 2; Length 224;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 EDSLIAA 427
Db 8 EDSLIAA 14

RESULT 15
PCT-US95-08565-16
; Sequence 16, Application PC/TUS9508565
```

```

/ GENERAL INFORMATION:
/ APPLICANT: Cashmore, Anthony R.
/ APPLICANT: Ahmad, Margaret
/ APPLICANT: Lin, Chentao
/ TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
/ TITLE OF INVENTION: Using the Same
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
/ STREET: One Liberty Place, 46th floor
/ CITY: Philadelphia
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19103
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/08565
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/272,255
/ FILING DATE: 08-JUL-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Leary Ph.D., Kathryn
/ REGISTRATION NUMBER: 36,317
/ REFERENCE/DOCKET NUMBER: UPN-1795
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (215) 568-3100
/ TELEFAX: (215) 568-3439
/ INFORMATION FOR SEQ ID NO: 16:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 224 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ PCT-US95-08565-16

```

```

Query Match 1.4%; Score 7; DB 5; Length 224;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 421 EDSLAA 427
Db 8 EDSLAA 14

```

Search completed: June 7, 2005, 17:11:48
Job time : 32 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:09:41 ; Search time 77.5 Seconds
(without alignments)
2370.369 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 510
Sequence: 1 MFEIHFVKSPPVVKYTEREI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1599520 seqs, 360203123 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	510	US-10-025-003-12	Sequence 12, Appl
2	510	100.0	510	US-10-718-952-12	Sequence 12, Appl
3	423	82.9	510	US-10-025-003-14	Sequence 14, Appl
4	423	82.9	510	US-10-025-003-16	Sequence 16, Appl
5	423	82.9	510	US-10-718-952-14	Sequence 14, Appl
6	423	82.9	510	US-10-718-952-16	Sequence 16, Appl
7	236	46.3	510	US-10-025-003-2	Sequence 2, Appl
8	236	46.3	510	US-10-025-003-10	Sequence 10, Appl
9	236	46.3	510	US-10-718-952-2	Sequence 2, Appl
10	236	46.3	510	US-10-718-952-10	Sequence 10, Appl
11	236	46.3	511	US-10-424-599-213009	Sequence 213009,
12	220	43.1	510	US-10-025-003-6	Sequence 6, Appl

13	220	43.1	510	US-10-718-952-6	Sequence 6, Appl
14	206	40.4	431	US-10-424-599-154863	Sequence 154863,
15	121	23.7	124	US-10-424-599-213004	Sequence 213004,
16	113	22.2	220	US-10-424-599-259439	Sequence 259439,
17	103	20.2	211	US-10-424-599-165505	Sequence 165505,
18	102	20.0	510	US-10-424-599-154864	Sequence 154864,
19	76	14.9	458	US-10-425-115-231843	Sequence 231843,
20	76	14.9	510	US-09-921-237-11	Sequence 11, Appl
21	76	14.9	510	US-09-921-330-11	Sequence 11, Appl
22	76	14.9	510	US-09-921-329-11	Sequence 11, Appl
23	76	14.9	510	US-10-767-701-46278	Sequence 46278, A
24	76	14.9	510	US-10-425-115-231852	Sequence 231852,
25	76	14.9	510	US-10-425-115-231857	Sequence 231857,
26	76	14.9	510	US-10-739-930-7635	Sequence 7635, Ap
27	76	14.9	516	US-10-425-114-62568	Sequence 62568, A
28	76	14.9	542	US-10-425-114-58674	Sequence 58674, A
29	67	13.1	332	US-10-425-114-42702	Sequence 42702, A
30	67	13.1	450	US-10-425-115-231845	Sequence 231845,
31	67	13.1	505	US-10-442-017-15	Sequence 15, Appl
32	67	13.1	510	US-09-727-628-2	Sequence 2, Appl
33	67	13.1	510	US-10-437-963-173252	Sequence 173252,
34	67	13.1	510	US-10-425-115-231850	Sequence 231850,
35	67	13.1	510	US-10-425-115-231857	Sequence 231857,
36	67	13.1	512	US-10-425-114-39621	Sequence 39621, A
37	67	13.1	512	US-10-425-114-53231	Sequence 53231, A
38	67	13.1	512	US-10-425-114-66216	Sequence 66216, A
39	55	10.8	509	US-10-437-963-187768	Sequence 187768,
40	49	9.6	250	US-10-424-599-154861	Sequence 154861,
41	47	9.2	78	US-10-424-599-213008	Sequence 213008,
42	44	8.6	394	US-10-363-829-458	Sequence 458, App
43	44	8.6	510	US-10-425-115-231853	Sequence 231853,
44	44	8.6	512	US-10-425-114-46916	Sequence 46916, A
45	42	8.2	108	US-10-425-115-231844	Sequence 231844,

ALIGNMENTS

RESULT 1
US-10-025-003-12
; Sequence 12, Application US/10025003
; Publication No. US2003007465A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-12
Query Match 100.0%; Score 510; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFEIHFVKSPPVVKYTEREI.....NIMRACVGLAPENNMLEYK 60
DB 1 MFEIHFVKSPPVVKYTEREI.....NIMRACVGLAPENNMLEYK 60
QY 61 KLGVWLVGNGNGSTLGVIANREDISWATKDKIQANVFGLTQASAIRVSGFQGBE 120


```
Db 61 KLGWLVGCGNGNGSTLTGCVIANREDISMAIKDKIQOANYGSLTQSAIRVSGFQGBE 120
Qy 121 IYAPFKSLPMPNPPDIYFGGWDISNMNLADAMAAKAFVDILOQKLRPYMESMPLPGI 180
Db 121 IYAPFKSLPMPNPPDIYFGGWDISNMNLADAMAAKAFVDILOQKLRPYMESMPLPGI 180
Qy 181 YPDPFIANOEERANNVIGKTKOEVOQIINDIKAFKATKVDKVVVMTANTERYSNLV 240
Db 181 YPDPFIANOEERANNVIGKTKOEVOQIINDIKAFKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAADRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIARNT 300
Db 241 VGLNDTMENLLAADRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIARNT 300
Qy 301 LIGGDFKSGQTKMSVLDPLVGAGIKPTISIVSNHGLGNDGMNLSAPQTRSKSISKS 360
Db 301 LIGGDFKSGQTKMSVLDPLVGAGIKPTISIVSNHGLGNDGMNLSAPQTRSKSISKS 360
Qy 361 NVDDMVNSNALIYEPGHPDHVVYIKVYPVYSGSKRAMDEYTSIFMGKXNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGHPDHVVYIKVYPVYSGSKRAMDEYTSIFMGKXNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSHFVATILSYLTAKPLVPPGTPVNN 480
Db 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSHFVATILSYLTAKPLVPPGTPVNN 480
Qy 481 ALSKORAMLENTMRACVGLAPENNNMLEYK 510
Db 481 ALSKORAMLENTMRACVGLAPENNNMLEYK 510
```

RESULT 2

```
US-10-718-952-12
; Sequence 12, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-12
```

```
Query Match 100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPELENKVSPPNKYETETIGSYVNTETTELHVENNGTYQWTVKRSVNYQFKNTNHP 60
Db 1 MPELENKVSPPNKYETETIGSYVNTETTELHVENNGTYQWTVKRSVNYQFKNTNHP 60
Qy 61 KLGWLVGCGNGNGSTLTGCVIANREDISMAIKDKIQOANYGSLTQSAIRVSGFQGBE 120
Db 61 KLGWLVGCGNGNGSTLTGCVIANREDISMAIKDKIQOANYGSLTQSAIRVSGFQGBE 120
Qy 121 IYAPFKSLPMPNPPDIYFGGWDISNMNLADAMAAKAFVDILOQKLRPYMESMPLPGI 180
Db 121 IYAPFKSLPMPNPPDIYFGGWDISNMNLADAMAAKAFVDILOQKLRPYMESMPLPGI 180
```

```
Qy 181 YPDPFIANOEERANNVIGKTKOEVOQIINDIKAFKATKVDKVVVMTANTERYSNLV 240
Db 181 YPDPFIANOEERANNVIGKTKOEVOQIINDIKAFKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAADRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIARNT 300
Db 241 VGLNDTMENLLAADRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIARNT 300
Qy 301 LIGGDFKSGQTKMSVLDPLVGAGIKPTISIVSNHGLGNDGMNLSAPQTRSKSISKS 360
Db 301 LIGGDFKSGQTKMSVLDPLVGAGIKPTISIVSNHGLGNDGMNLSAPQTRSKSISKS 360
Qy 361 NVDDMVNSNALIYEPGHPDHVVYIKVYPVYSGSKRAMDEYTSIFMGKXNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGHPDHVVYIKVYPVYSGSKRAMDEYTSIFMGKXNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSHFVATILSYLTAKPLVPPGTPVNN 480
Db 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSHFVATILSYLTAKPLVPPGTPVNN 480
Qy 481 ALSKORAMLENTMRACVGLAPENNNMLEYK 510
Db 481 ALSKORAMLENTMRACVGLAPENNNMLEYK 510
```

RESULT 3

```
US-10-025-003-14
; Sequence 14, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-14
```

```
Query Match 82.9%; Score 423; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 88 ISMATKDKIQOANYGSLTQSAIRVSGFQGBEIIYAPFKSLPMPNPPDIYFGGWDISNM 147
Db 88 ISMATKDKIQOANYGSLTQSAIRVSGFQGBEIIYAPFKSLPMPNPPDIYFGGWDISNM 147
Qy 148 NLADAMAAKAFVDILOQKLRPYMESMPLPGIYDPDFIAANOEERANNVIGKTKOEVOQ 207
Db 148 NLADAMAAKAFVDILOQKLRPYMESMPLPGIYDPDFIAANOEERANNVIGKTKOEVOQ 207
Qy 208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGLNDTMENLLAADRNEAISPSTLY 267
Db 208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGLNDTMENLLAADRNEAISPSTLY 267
Qy 268 AIACWENVPFINSPOPTFVGLIDLAIARNTLIGGDFKSGQTKMSVLDPLVGAGI 327
Db 268 AIACWENVPFINSPOPTFVGLIDLAIARNTLIGGDFKSGQTKMSVLDPLVGAGI 327
Qy 328 KPTISIVSNHGLGNDGMNLSAPQTRSKSISKSNNVDDMVNSNALIYEPGHPDHVVYIK 387
```

```

Db      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
Qy      448 ENEGKFHSHFPAVATILSYLTAKAPLVPGTTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTAKAPLVPGTTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
Db      508 EYK 510
      508 EYK 510
      508 EYK 510

RESULT 4
US-10-025-003-16
; Sequence 16, Application US/10025003
; Publication No. US20030074665A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-16

Query Match      82.9%; Score 423; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 ISMATKDKIQOANYFGSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNM 147
      88 ISMATKDKIQOANYFGSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNM 147
Db      148 NIADMAARAKVFDIDLQKQLRPYMESWVPLPGIYDPDFIAANOEBRANNVIKGTQOEQVQ 207
      148 NIADMAARAKVFDIDLQKQLRPYMESWVPLPGIYDPDFIAANOEBRANNVIKGTQOEQVQ 207
Qy      208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGNDTMENLLAADVREAEISPSSTLY 267
      208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGNDTMENLLAADVREAEISPSSTLY 267
Db      208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGNDTMENLLAADVREAEISPSSTLY 267
      208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGNDTMENLLAADVREAEISPSSTLY 267
Qy      268 AIACVMEVNPFIINGSPQNTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLDVFLVGAGI 327
      268 AIACVMEVNPFIINGSPQNTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLDVFLVGAGI 327
Db      268 AIACVMEVNPFIINGSPQNTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLDVFLVGAGI 327
      268 AIACVMEVNPFIINGSPQNTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLDVFLVGAGI 327
Qy      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
Db      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
Qy      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
Db      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
Qy      448 ENEGKFHSHFPAVATILSYLTAKAPLVPGTTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTAKAPLVPGTTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTAKAPLVPGTTPVNNALSKORAMLENIMRACVGLAPENNMTL 507

```

```

Qy      508 EYK 510
      508 EYK 510
      508 EYK 510

RESULT 5
US-10-718-952-14
; Sequence 14, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-14

Query Match      82.9%; Score 423; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 ISMATKDKIQOANYFGSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNM 147
      88 ISMATKDKIQOANYFGSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNM 147
Db      148 NIADMAARAKVFDIDLQKQLRPYMESWVPLPGIYDPDFIAANOEBRANNVIKGTQOEQVQ 207
      148 NIADMAARAKVFDIDLQKQLRPYMESWVPLPGIYDPDFIAANOEBRANNVIKGTQOEQVQ 207
Qy      208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGNDTMENLLAADVREAEISPSSTLY 267
      208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGNDTMENLLAADVREAEISPSSTLY 267
Db      208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGNDTMENLLAADVREAEISPSSTLY 267
      208 QIIKQIKAFKATKVDKVVVMTANTERYSNLVGNDTMENLLAADVREAEISPSSTLY 267
Qy      268 AIACVMEVNPFIINGSPQNTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLDVFLVGAGI 327
      268 AIACVMEVNPFIINGSPQNTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLDVFLVGAGI 327
Db      268 AIACVMEVNPFIINGSPQNTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLDVFLVGAGI 327
      268 AIACVMEVNPFIINGSPQNTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLDVFLVGAGI 327
Qy      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
Db      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
      328 KPTSIYSVNHGNDGNMISAPQTFRSKISKSNNVDDMNSNALIYEGEHPDHVVYK 387
Qy      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
Db      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLAELSTRIOFKA 447
Qy      448 ENEGKFHSHFPAVATILSYLTAKAPLVPGTTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTAKAPLVPGTTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTAKAPLVPGTTPVNNALSKORAMLENIMRACVGLAPENNMTL 507

```

GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 16
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-718-952-16

Query Match 82.9%; Score 423; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ISMATKDKIQOANFGSLTQASAIRVSGFGEIYAPFKSLPMVNPDDIVFGMDISM 147
DB 88 ISMATKDKIQOANFGSLTQASAIRVSGFGEIYAPFKSLPMVNPDDIVFGMDISM 147
QY 148 NLADAMARAKVPDIDLQKLRPYMESMVLPGIYBDFIAANOEBRANNVIKGTQEOVQ 207
DB 148 NLADAMARAKVPDIDLQKLRPYMESMVLPGIYBDFIAANOEBRANNVIKGTQEOVQ 207
QY 208 QIKKIKAKKATKVDKVVVLTANTERSNLVGNLDMENLLAANDNEAISPSTLY 267
DB 208 QIKKIKAKKATKVDKVVVLTANTERSNLVGNLDMENLLAANDNEAISPSTLY 267
QY 268 AIAQWENVPFINGSPQNTFVPGGLIDLAIAANTLIGDDPKSGQTKMSVLDVVGAGI 327
DB 268 AIAQWENVPFINGSPQNTFVPGGLIDLAIAANTLIGDDPKSGQTKMSVLDVVGAGI 327
QY 328 KPTSIYSVNLGNDGMNLSAQTFERSKISKSNVDDMVNSNAILYEPGEHPDHVVVK 387
DB 328 KPTSIYSVNLGNDGMNLSAQTFERSKISKSNVDDMVNSNAILYEPGEHPDHVVVK 387
QY 388 YVPYGDSTRANDVETSEIFMGSKNTIYLAHNTCEBSLLAAILDLVLAELSTRIOFKA 447
DB 388 YVPYGDSTRANDVETSEIFMGSKNTIYLAHNTCEBSLLAAILDLVLAELSTRIOFKA 447
QY 448 ENEGKFSHPVATILSYLTAKALVPGTPPVNALSOKRAMLENIMRACVGLAPENMIL 507
DB 448 ENEGKFSHPVATILSYLTAKALVPGTPPVNALSOKRAMLENIMRACVGLAPENMIL 507
QY 508 EYK 510
DB 508 EYK 510

RESULT 7
US-10-025-003-2
Sequence 2, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07

PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 2
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-2

Query Match 46.3%; Score 236; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.1e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYBDFIAANOEBRANNVIKGTQEOVQOIKIKAFKATKVDKVVVLTANTER 235
DB 176 PLPGIYBDFIAANOEBRANNVIKGTQEOVQOIKIKAFKATKVDKVVVLTANTER 235
QY 236 YSNLVGNDTMENLLAANDNEAISPSTLYAIAQWENVPFINGSPQNTFVPGGLIDLA 295
DB 236 YSNLVGNDTMENLLAANDNEAISPSTLYAIAQWENVPFINGSPQNTFVPGGLIDLA 295
QY 296 IARNTLIGDDPKSGQTKMSVLDVVGAGIKPTSIYSVNLGNDGMNLSAQTFERSK 355
DB 296 IARNTLIGDDPKSGQTKMSVLDVVGAGIKPTSIYSVNLGNDGMNLSAQTFERSK 355
QY 356 EISKSNVDDMVNSNAILYEPGEHPDHVVVKYVPYGDSTRANDVETSEIFMGSK 411
DB 356 EISKSNVDDMVNSNAILYEPGEHPDHVVVKYVPYGDSTRANDVETSEIFMGSK 411

RESULT 8
US-10-025-003-10
Sequence 10, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 10
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-10

Query Match 46.3%; Score 236; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.1e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYBDFIAANOEBRANNVIKGTQEOVQOIKIKAFKATKVDKVVVLTANTER 235
DB 176 PLPGIYBDFIAANOEBRANNVIKGTQEOVQOIKIKAFKATKVDKVVVLTANTER 235
QY 236 YSNLVGNDTMENLLAANDNEAISPSTLYAIAQWENVPFINGSPQNTFVPGGLIDLA 295
DB 236 YSNLVGNDTMENLLAANDNEAISPSTLYAIAQWENVPFINGSPQNTFVPGGLIDLA 295
QY 296 IARNTLIGDDPKSGQTKMSVLDVVGAGIKPTSIYSVNLGNDGMNLSAQTFERSK 355
DB 296 IARNTLIGDDPKSGQTKMSVLDVVGAGIKPTSIYSVNLGNDGMNLSAQTFERSK 355

Db 296 IARNTLIGDDDFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDDGNLSAPQTFRSK 355
QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSEIFMGSK 411
Db 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSEIFMGSK 411

RESULT 9
US-10-718-952-2

Sequence 2, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Stetlic, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-718-952-2

Query Match 46.3%; Score 236; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.1e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDDPFIANOEBRANVYIKGTQKQOVQIIKDIAFKFEATKVDKVVVLTANTR 235
Db 176 PLPGIYDDPFIANOEBRANVYIKGTQKQOVQIIKDIAFKFEATKVDKVVVLTANTR 235
QY 236 YSNLVGGLNDTMENTLLAAVDRNEAISPSTLYAIACWENVPFINGSPONTFVPGILIDA 295
Db 236 YSNLVGGLNDTMENTLLAAVDRNEAISPSTLYAIACWENVPFINGSPONTFVPGILIDA 295
QY 236 IARNTLIGDDDFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDDGNLSAPQTFRSK 355
Db 236 IARNTLIGDDDFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDDGNLSAPQTFRSK 355
QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSEIFMGSK 411
Db 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSEIFMGSK 411

RESULT 10
US-10-718-952-10

Sequence 10, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Stetlic, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97
SEQ ID NO 10
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-718-952-10

Query Match 46.3%; Score 236; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.1e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDDPFIANOEBRANVYIKGTQKQOVQIIKDIAFKFEATKVDKVVVLTANTR 235
Db 176 PLPGIYDDPFIANOEBRANVYIKGTQKQOVQIIKDIAFKFEATKVDKVVVLTANTR 235
QY 236 YSNLVGGLNDTMENTLLAAVDRNEAISPSTLYAIACWENVPFINGSPONTFVPGILIDA 295
Db 236 YSNLVGGLNDTMENTLLAAVDRNEAISPSTLYAIACWENVPFINGSPONTFVPGILIDA 295
QY 236 IARNTLIGDDDFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDDGNLSAPQTFRSK 355
Db 236 IARNTLIGDDDFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDDGNLSAPQTFRSK 355
QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSEIFMGSK 411
Db 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSEIFMGSK 411

RESULT 11

US-10-424-599-213009
Sequence 213009, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 213009
LENGTH: 511
TYPE: PRT
ORGANISM: Glycine max
NAME/KEY: unsure
LOCATION: (1)..(511)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.pep
US-10-424-599-213009

Query Match 46.3%; Score 236; DB 15; Length 511;
Best Local Similarity 100.0%; Pred. No. 7.2e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDDPFIANOEBRANVYIKGTQKQOVQIIKDIAFKFEATKVDKVVVLTANTR 235
Db 176 PLPGIYDDPFIANOEBRANVYIKGTQKQOVQIIKDIAFKFEATKVDKVVVLTANTR 236
QY 236 YSNLVGGLNDTMENTLLAAVDRNEAISPSTLYAIACWENVPFINGSPONTFVPGILIDA 295
Db 236 YSNLVGGLNDTMENTLLAAVDRNEAISPSTLYAIACWENVPFINGSPONTFVPGILIDA 296
QY 236 IARNTLIGDDDFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDDGNLSAPQTFRSK 355
Db 236 IARNTLIGDDDFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDDGNLSAPQTFRSK 356
QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSEIFMGSK 411
Db 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSEIFMGSK 411

Db 357 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRAMEDTSEIFMGSK 412

RESULT 12

US-10-025-003-6

Sequence 6, Application US/10025003

Publication No. US20030074685A1

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

APPLICANT: Streitz, Leon

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-C

CURRENT APPLICATION NUMBER: US/10/025, 003

CURRENT FILING DATE: 2002-05-07

PRIOR APPLICATION NUMBER: 08/835, 751

PRIOR FILING DATE: APRIL 8, 1997

PRIOR APPLICATION NUMBER: PCT/US98/06822

PRIOR FILING DATE: APRIL 7, 1998

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97

SEQ ID NO 6

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-10-025-003-6

Query Match 43.1%; Score 220; DB 14; Length 510;

Best Local Similarity 100.0%; Pred. No. 5.3e-209;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 PLPGIYDPDFIAANOEBRANNVIKGTQKQVQOIIKDIKAFKATKYVDKVVVLTANTER 235

Db 176 PLPGIYDPDFIAANOEBRANNVIKGTQKQVQOIIKDIKAFKATKYVDKVVVLTANTER 235

Qy 236 YSNLVGNDTMENTLLAANDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDIA 295

Db 236 YSNLVGNDTMENTLLAANDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDIA 295

Qy 236 IARNTLIGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Db 236 IARNTLIGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Qy 236 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGD 395

Db 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGD 395

RESULT 13

US-10-718-952-6

Sequence 6, Application US/10718952

Publication No. US20040128713A1

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

APPLICANT: Streitz, Leon

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-C

CURRENT APPLICATION NUMBER: US/10/718, 952

CURRENT FILING DATE: 2003-11-21

PRIOR APPLICATION NUMBER: 08/835, 751

PRIOR FILING DATE: APRIL 8, 1997

PRIOR APPLICATION NUMBER: PCT/US98/06822

PRIOR FILING DATE: APRIL 7, 1998

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97

SEQ ID NO 6

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-10-718-952-6

Query Match 43.1%; Score 220; DB 16; Length 510;

Best Local Similarity 100.0%; Pred. No. 5.3e-209;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 PLPGIYDPDFIAANOEBRANNVIKGTQKQVQOIIKDIKAFKATKYVDKVVVLTANTER 235

Db 176 PLPGIYDPDFIAANOEBRANNVIKGTQKQVQOIIKDIKAFKATKYVDKVVVLTANTER 235

Qy 236 YSNLVGNDTMENTLLAANDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDIA 295

Db 236 YSNLVGNDTMENTLLAANDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDIA 295

Qy 236 IARNTLIGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Db 236 IARNTLIGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Qy 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGD 395

Db 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGD 395

RESULT 14

US-10-424-599-154863

Sequence 154863, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J

APPLICANT: Kovalic, David K

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424, 599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 154863

LENGTH: 431

TYPE: PRT

ORGANISM: Glycine max

FEATURES:

OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1.pcp

US-10-424-599-154863

Query Match 40.4%; Score 206; DB 15; Length 431;

Best Local Similarity 99.7%; Pred. No. 3.5e-195;

Matches 306; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 88 ISMATKDKIQOANYFGSLTOASAIRVGSFOGEBIYAPFKSLIPMVNPDIYFGMDISNM 147

Db 88 ISMATKDKIQOANYFGSLTOASAIRVGSFOGEBIYAPFKSLIPMVNPDIYFGMDISNM 147

Qy 148 NLADAMARAKVFDIDLOKOLRPYMESVPLPGIYDPDFIAANOEBRANNVIKGTQKQVQ 207

Db 148 NLADAMARAKVFDIDLOKOLRPYMESVPLPGIYDPDFIAANOEBRANNVIKGTQKQVQ 207

Qy 208 QIHKDKAFKATKYVDKVVVLTANTERYSNLVGLNDTMENTLLAANDRNEAEISPTLY 267

Db 208 QIHKDKAFKATKYVDKVVVLTANTERYSNLVGLNDTMENTLLAANDRNEAEISPTLY 267

Qy 268 AIACWENVPFINGSPONTFVPGILDIAIARNTLIGDDPKSGQTKMSVLVDPLVGAGI 327

Db 268 AIACWENVPFINGSPONTFVPGILDIAIARNTLIGDDPKSGQTKMSVLVDPLVGAGI 327

Qy 328 KPTSIYSYNHLGNDGMNLSAPQTFRSKEISKSNVDDMVNSNALIYEPGEHPDHVVVIK 387

Db 328 KPTSIYSYNHLGNDGMNLSAPQTFRSKEISKSNVDDMVNSNALIYEPGEHPDHVVVIK 387

Qy 388 YVPYVGD 394

Db 388 YVPVVD 394

RESULT 15

US-10-424-599-213004

/ Sequence 213004, Application US/10424599

/ Publication No. US20040031072A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa Thomas J

/ APPLICANT: Kovalic David K

/ APPLICANT: Zhou Yihua

/ APPLICANT: Cao Yongwei

/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

/ FILE REFERENCE: 38-21(53223)B

/ CURRENT APPLICATION NUMBER: US/10/424,599

/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 285684

/ SEQ ID NO 213004

/ LENGTH: 124

/ TYPE: PRT

/ ORGANISM: Glycine max

/ FEATURE:

/ OTHER INFORMATION: Clone ID: PAT_MRT3847_34368C.1.pep

US-10-424-599-213004

Query Match 23.7%; Score 121; DB 15; Length 124;

Best Local Similarity 100.0%; Pred. No. 2.2e-11;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	390	PVVGDSKRAMDEYTSRIFMGKNTIVLHNTCEDSLAPITIDLVYLAELSTRIOPKAEN	449
DB	4	PVVGDSKRAMDEYTSRIFMGKNTIVLHNTCEDSLAPITIDLVYLAELSTRIOPKAEN	63
QY	450	BGKFSHFHVATILSYLTAKPLVPPTPVVNALSKORAMLENIMRACVGLAPENNMILEY	509
DB	64	BGKFSHFHVATILSYLTAKPLVPPTPVVNALSKORAMLENIMRACVGLAPENNMILEY	123
QY	510	K 510	
DB	124	K 124	

Search Completed: June 7, 2005, 17:25:34
Job time : 78.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 15:12:39 ; Search time 86 Seconds
(without alignments) 3036.749 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MFENFKVESPNVKTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2593	98.6	510	2 094C02	094C02 glycine max
2	2580	98.1	510	2 094K03	094K03 glycine max
3	2475	94.1	510	1 INO1_TOBAC	091W96 nicotiana t
4	2470	93.9	510	1 INO1_NICPA	0958V4 nicotiana p
5	2452	93.2	510	1 INO1_SESIN	095FV1 sesamum ind
6	2400	91.2	510	1 INO1_WHEBAT	0947U0 triticum ae
7	2399	91.2	510	2 0944G3	0944G3 suadae sals
8	2395	91.0	512	1 INO1_MESCR	040271 mesembryant
9	2384	90.6	510	1 INO2_ARATH	038862 arabidopsi
10	2376	90.3	510	1 INO1_BRANA	096348 brassica na
11	2376	90.3	510	1 INO3_ARATH	091X12 arabidopsi
12	2373.5	90.2	509	2 09AR12	09AR12 avicennia m
13	2361	89.7	510	2 07XJCO	07XJCO xerophyta v
14	2358	89.6	510	1 INO1_SPIPO	042803 spiridocela p
15	2354	89.5	510	1 INO1_MAIZE	095FPX zea mays (m
16	2354	89.5	510	2 09AV99	09AV99 avena sativ
17	2349.5	89.3	511	1 INO1_ARATH	094801 arabidopsi
18	2345	89.1	510	1 INO1_ORYSA	064437 oryza sativ
19	2338	88.9	510	1 INO1_KORYU	065195 hordeum vul
20	2337	88.8	472	2 094G22	094G22 phaseolus v
21	2328.5	88.5	511	1 INO1_PITPA	041107 phaseolus v
22	2304.5	87.6	507	1 INO1_CITPA	042802 citrus para
23	2221.5	84.4	509	2 08SSN2	08SSN2 oryza sativ
24	2207	76.3	443	2 08H1B7	08H1B7 lolium pere
25	1919.5	73.0	409	2 084SA1	084SA1 aster tipo
26	1883	71.6	512	2 07XZB6	07XZB6 porteresia
27	1728	65.7	364	2 09AWG8	09AWG8 actinidia a
28	1722.5	65.5	563	2 07ZXVO	07ZXVO xenopus lae
29	1716.5	65.2	560	2 06DDT1	06DDT1 xenopus lae
30	1616	61.4	558	2 09NPH2	09NPH2 homo sapien
31	1616	61.4	558	2 09NVW7	09NVW7 homo sapien

32	1611	61.2	558	2 09H2Y2	09H2Y2 homo sapien
33	1606	61.0	537	2 087JUS	087JUS neurospora
34	1606	61.0	537	2 09JHU9	09JHU9 m myo-inosi
35	1603	60.9	558	2 06NXT5	06NXT5 homo sapien
36	1602	60.9	565	1 INO1_DROME	097477 drosophila
37	1586.5	60.3	561	2 07PZB9	07PZB9 anophelis g
38	1566.5	59.5	526	2 07Z525	07Z525 homo sapien
39	1525	58.0	526	2 095PT4	095PT4 leishmania
40	1509	57.4	526	2 09N9B7	09N9B7 leishmania
41	1476	56.1	530	2 06CF53	06CF53 yarrowia li
42	1472	55.9	508	2 P90626	P90626 entamoeba h
43	1469.5	55.9	527	2 000815	000815 leishmania
44	1459.5	55.5	520	2 06BJ15	06BJ15 debaryomyce
45	1450	55.1	525	2 074247	074247 plichia past

ALIGNMENTS

RESULT 1					
ID	094C02	PRELIMINARY;	PRT;	510 AA.	
AC	094C02;				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)				
DE	Myo-inositol-1-phosphate synthase (EC 5.5.1.4).				
OS	Glycine max (Soybean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
NCBI	NCBI_TaxID=3847;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RA	Hitz W.D., Carlson T.J., Kerr P., Sebastian S.;				
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY038802; AAK72098.1; --				
DR	HSSP; P18861; 1PIH.				
DR	GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.				
DR	GO; GO:0016853; F:isomerase activity; IEA.				
DR	GO; GO:0006021; P:myo-inositol biosynthesis; IEA.				
DR	GO; GO:0008654; P:phospholipid biosynthesis; IEA.				
DR	InterPro: IPR002587; Inos-1-P_synth.				
PFam	PF01658; Inos-1-P_synth; 1.				
KW	Isomerase.				
SC	SEQUENCE 510 AA; 56475 MW; A1B4C77F1643918E CRC64;				
Query Match 98.6%; Score 2593; DB 2; Length 510;					
Best Local Similarity 98.4%; Pred. No. 2.6e-162;					
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;					
QY	1	MFENFKVESPNVKTETETI	QV	1	MFENFKVESPNVKTETETI
DB	1	MFENFKVECNVKTETETI	QV	1	MFENFKVECNVKTETETI
QY	61	KLGVMLVGMGNNSTLTGVI	QV	61	KLGVMLVGMGNNSTLTGVI
DB	61	KLGVMLVGMGNNSTLTGVI	QV	61	KLGVMLVGMGNNSTLTGVI
QY	121	IYAPFKSLPMVNPDDIVFG	QV	121	IYAPFKSLPMVNPDDIVFG
DB	121	IYAPFKSLPMVNPDDIVFG	QV	121	IYAPFKSLPMVNPDDIVFG
QY	181	YDPDFIANGBERANNVIG	QV	181	YDPDFIANGBERANNVIG
DB	181	YDPDFIANGBERANNVIG	QV	181	YDPDFIANGBERANNVIG
QY	241	VGNDTMENTLAAVDRNEA	QV	241	VGNDTMENTLAAVDRNEA
DB	241	VGNDTMENTLAAVDRNEA	QV	241	VGNDTMENTLAAVDRNEA
QY	301	LIGDDPFGSGOTRKSVLV	QV	301	LIGDDPFGSGOTRKSVLV
DB	301	LIGDDPFGSGOTRKSVLV	QV	301	LIGDDPFGSGOTRKSVLV

```

Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSNHLGNNDGNMLSAPOFRSKSISKS 360
Qy 361 NVVDMDVNSNAIIVPEGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEIFMGKNITVLHNTC 420
Db 361 NVVDMDVNSNAIIVPEGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEIFMGKNITVLHNTC 420
Qy 421 EDLSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLPVPGTPTPVN 480
Db 421 EDLSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLPVPGTPTPVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 2

```

Q94KU3 PRELIMINARY; PRT; 510 AA.
ID Q94KU3
AC Q94KU3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Myo-inositol-3-phosphate synthase.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OC NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21196082; PubMed=11299373; DOI=10.1104/PP.125.4.1941;
RA Hegeman C.E., Good L.L., Grabau E.A.;
RT "Expression of D-myo-inositol-3-phosphate synthase in soybean.
RT Implications for phytic acid biosynthesis.";
RL Plant Physiol. 125:1941-1948 (2001).
DR EMBL; AF293970; AAK49896.1; -.
DR HSSP; P11986; 1PIH.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth; 1.
SQ SEQUENCE 510 AA; 56506 MW; DE4F3DD7DC6F370 CRC64;

```

```

Query Match 98.1%; Score 2580; DB 2; Length 510;
Best Local Similarity 98.0%; Pred. No. 1.9e-161;
Matches 500; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 1 MFIEFKVESPNVKTETETIOSVYNYETTELVEHNRNGTYQWIVKPKSVNYQFKTNTHP 60
Db 1 MFIEFKVECNVKTETETIOSVYNYETTELVEHNRNGTYQWIVKPKSVNYQFKTNTHP 60
Qy 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVSGFOGEE 120
Db 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVSGFOGEE 120
Qy 121 IYAPFKSLPWNPDIVFGGWDISNNMLADAMAAKAFVDIDLQKLRPYMESWVPLPGI 180
Db 121 IYAPFKSLPWNPDIVFGGWDISNNMLADAMAAKAFVDIDLQKLRPYMESWVPLPGI 180
Qy 181 YDPPFIANOBERANVVKGTQKQOVQOIIDKIKAFKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPPFIANOBERANVVKGTQKQOVQOIIDKIKAFKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPFINSQPTFVGLIDLAIANT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPFINSQPTFVGLIDLAIANT 300
Qy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSNHLGNNDGNMLSAPOFRSKSISKS 360
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSNHLGNNDGNMLSAPOFRSKSISKS 360

```

```

Qy 361 NVVDMDVNSNAIIVPEGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEIFMGKNITVLHNTC 420
Db 361 NVVDMDVNSNAIIVPEGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEIFMGKNITVLHNTC 420
Qy 421 EDLSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLPVPGTPTPVN 480
Db 421 EDLSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLPVPGTPTPVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 3

```

INOL_TOBAC STANDARD; PRT; 510 AA.
ID INOL_TOBAC
AC Q9LW96;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (MI-1-P synthase) (IPS).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OC NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=leaf;
RX MEDLINE=20399434; PubMed=10945337;
RA Hara K., Yagi M., Koizumi N., Kusano T., Sano H.;
RT "Screening of wound-responsive genes identifies an immediate-early
RT expressed gene encoding a highly charged protein in mechanically
RT wounded tobacco plants.";
RL Plant Cell Physiol. 41:684-691 (2000).
RL Plant Cell Physiol. 41:684-691 (2000).
CC -1 CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-
CC phosphate.
CC -1 COPACTOR: NAD (By similarity).
CC -1 PATHWAY: Inositol biosynthesis.
CC -1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1 SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.

```

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

```

DR EMBL; AB009881; BAA95788.1; -.
DR HSSP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth. 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56369 MW; 4BA8FD5ADB6D4D CRC64;

```

```

Query Match 94.1%; Score 2475; DB 1; Length 510;
Best Local Similarity 92.7%; Pred. No. 1.5e-154;
Matches 473; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy 1 MFIEFKVESPNVKTETETIOSVYNYETTELVEHNRNGTYQWIVKPKSVNYQFKTNTHP 60
Db 1 MFIEFKVESPNVKTETETIOSVYNYETTELVEHNRNGTYQWIVKPKSVNYQFKTNTHP 60
Qy 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVSGFOGEE 120
Db 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVSGFOGEE 120
Qy 121 IYAPFKSLPWNPDIVFGGWDISNNMLADAMAAKAFVDIDLQKLRPYMESWVPLPGI 180
Db 121 IYAPFKSLPWNPDIVFGGWDISNNMLADAMAAKAFVDIDLQKLRPYMESWVPLPGI 180

```

Db	121	YAPFESKLLPMWNPDVVYFGWDISGMNLADAMAAKXVPDIDLGKLRPYMESMPLPGI	180
Qy	181	YDPDPIAANQGEERANNVTKGTQOEVOQIHKIKAFKEATKYDKRVVLTANTERYSNLV	240
Db	181	YDPDPIAANQGEERANNVTKGTQOEQIDQIIKIDREEKEKXKVDKVVVLTANTERSNVV	240
Qy	241	VGLNNTMENLLAAVDNRNEAISPSTLYAACMENVPFINSGSPQNTFVPGCLIDLAARNT	300
Db	241	VGLNNTMENLFPASVBRNEAEISPSTLYAACILENVPFINSGSPQNTFVPGCLIDLAARNT	300
Qy	301	LIGGDDFSGQTKMKSVLVDPLVAGIKIKPSTSVSYNHLGNDDGMLSAPOTFRSKTISKS	360
Db	301	LIGGDDFSGQTKMKSVLVDPLVAGIKIKPSTSVSYNHLGNDDGMLSAQOTFRSKTISKS	360
Qy	361	NVVDVMVNSNALYEPGEHPDHVVYIKVYPYVGDSCRAMDEYTSSEIFMGQKNTIVLHNTC	420
Db	361	NVVDVMVNSNALYEPGEHPDHVVYIKVYPYVGDSCRAMDEYTSSEIFMGQKNTIVLHNTC	420
Qy	421	EDSLLAAPILDDVLLVLAELSTRIOQKAENEGKRFHSHPVATILSYTKAPLYPFGTTPVVN	480
Db	421	EDSLLAAPILDDVLLVLAELSTRIOQKAEEGKRFHSHPVATILSYTKAPLYPFGTTPVVN	480
Qy	481	ALSKORAMLENTMRACVGLAPENNMLLEVK	510
Db	481	ALSKORAMLENTLRACVGLAPENNMLLEVK	510

RESULT 4	
ID	STANDARD;
IN01_NICPA	
IN01_NICPA	PRT; 510 AA.

DT 16-OCR-2001 (Rel. 40, Last sequence update)
 PR 05-Jun-2004 (Rel. 44, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
 synthase) (M1-1-P synthase) (1PS).
 DE Name=INPS1;
 OS Nicotiana paniculata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=62141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hashimoto A., Yamada S., Komori T.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-
 CC phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
 CC family.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).

DR InterPro: IPR002587; Inos-1-P synth.
Pfam: PF01658; Inos-1-P synth; 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis
SQ SEQUENCE 510 AA; 5638 MW; 41B81C7A267666 CRC64;

Query Match	93.9%;	Score 2470;	DB 1;	Length 510;
Best Local Similarity	92.4%;	Pred. No. 3.2e-154;		
Matches 471;	Conservative 21;	Mismatches 18;	Indels 0;	Gaps 0;

Qy	1	MEIEBKVESPVVKYTEHIEIOSVNAVYETTELVEHNRNGYOMIVKRSXNYOGKNTMHP	60
Db	1	MEIENFKVESPVVKTESIEHSVVOYQTELLVHDEKNIGYQMTVKPKTYAEKTDVHP	60
Qy	61	KLGVNLVGMGNGNGSTLTCGVYANRBDISMTATKDKIQOANYGSGSLQOAAIRVSGFQEE	120
Db	61	KLGVNLVGMGNGNGSTLTCGVYANRBDISMTATKDKIQOANYGSGSLQOASTIRVSGNGEE	120
Qy	121	IYAPFKSLTPVNPDPDIVFGWDISNMNLADAMARAKVPDIDLQOKLRPYMESVPLPGI	180
Db	121	IYAPFKSLTPVNPDPDIVFGWDISNMNLADAMARAKVPDIDLQOKLRPYMESVPLPGI	180
Qy	181	YDPDEITAAOGERANNVIKGTQOEVOOQIKOIKAKFEKTKYDKVYVLTANTERISNLV	240
Db	181	YDPDEITAAOGRANNVIKGTKEQIDQIKOIREKFEKTKYDKVYVLTANTERYSNV	240
Qy	241	VGINDTMENLAAVDRNEAETSPTLYIAACMENVPFINGSPONTFVGLDILAIABNT	300
Db	241	VGINDTMENLFPASVDRNEAETSPTLYIACLLENVPFINGSPONTFVPLDILAIKNT	300
Qy	301	LIGGDDFSGQTKMSVLVDFLVGAGIKPETSIVSYNHLGNNDGMNLSAPOTFSKEISKS	360
Db	301	LIGGDDFSGQTKMSVLVDFLVGAGIKPETSIVSYNHLGNNDGMNLSAQPTFSKEISKS	360
Qy	361	NVDDMVNSNALYEPGEHPDHVVIKYVPYVGSKRADDEYTSBEIFMGSKNTIVLANTC	420
Db	361	NVDDMVNSNALYEPGEHPDHVVIKYVPYVGSKRADDEYTSBEIFMGSKNTIVLANTC	420
Qy	421	EDSLLAAPILIDVLLAELSTRIOKFAENEGKFHSFHPATILSYTKAPLVPFGTPVNV	480
Db	421	EDSLLAAPILIDVLLAELSTRIOKFAEBEGKFHSFHPATILSYTKAPLVPFGTPVNV	480
Qy	481	ALSKORAMLENIMRACVGLAPENNMLLEKX	510
Db	481	ALSKORAMLENIMRACVGLAPENNMLILEKX	510

RESULT 5	
INO1_SESIN	
ID	INO1_SESIN
	STANDARD;
	PRT; 510 AA

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (M1-1-P synthase) (IP3).
OS *Sesamum indicum* (Oriental sesame) (gingelly).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Pedaliaceae; *Sesamum*.
OX NCBI_TaxID=4182;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RA jin U.-H., Chung C.-H.;
RT "Characterization and functional analysis of a myo-inositol 1-
RT phosphate synthase cDNA from sesame (*Sesamum indicum* L.) seeds.";
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-
CC phosphate.
CC -1 COFACTOR: NAD (By similarity).
CC -1 PATHWAY: Inositol biosynthesis.
CC -1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1 SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.

CC This SWISSPORT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its use
CC by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/announce/
CC or_send_an_email_to_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Amaranthaceae; Suaeda.
 OC NCBI_TaxID=126914;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang J.P., Mang P.P., Sun Y.F., Zhao Y.X., Zhang H.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBS databases.
 DR EMBL; AF433879; AAL28131.1; --
 DR HSSP; P11986; 1PIK.
 DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR SEQUENCE 510 AA; 5665 MW; 6C31006D2B1C508A CRC64;
 Query Match 91.2%; Score 2399; DB 2; Length 510;
 Best Local Similarity 90.4%; Pred. No. 1.5e-149;
 Matches 462; Conservative 22; Mismatches 25; Indels 2; Gaps 2;
 QY 1 MFIEHFVSPVVKYETETETIOSVNYETTELVEHNRNGT-YQWIVKPKSVNYQFKTTHV 59
 DB 1 MFIEHFVSPVVKYETETETIOSVNYETTELVEHNRKSGVQWIVKPKSVNYQFKTTHV 60
 QY 60 PKLGMLVKGNGNSTLTGVIANREDISWATKXIQOANFGSLTQASAIRVSGFOE 119
 DB 61 PKLGMLVKGNGNSTLTGVIANREGISWATKXIQOANFGSLTQASAIRVSGFNGE 120
 QY 120 EIIYAPFKSLPWNPPDDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESHVPLPG 179
 DB 121 EIIYAPFKSLPWNPPDDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESHVPLPG 180
 QY 180 IYDPDFIAANOEBRANNVIKGTQOEVOQOIIKDIKAFKEATKVDKVVVLTANTERYSNL 239
 DB 181 IYDPDFIAANOEBRANNVIKGTQOEVOQOIIKDIKAFKEATKVDKVVVLTANTERYSDV 240
 QY 240 VVGANDTMENLLAVDRNEAISESTLYAIACWENVPFINGSPONTFVPGILDALAIRN 299
 DB 241 VVGANDTMENLLAVDRNEAISESTLYAIACWENVPFINGSPONTFVPGILDALAIRN 300
 QY 300 TLIGGDDPKSGGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDSAPOTFRSKEISK 359
 DB 301 CLIGGDDPKSGGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDSAPOTFRSKEISK 360
 QY 360 SNVVDVNVSNALILYEPGEHPDHVVVVKVPPVGDGSKAMDEYTSIEIFMGKNTIVLHNT 419
 DB 361 SNVVDVNVSNALILYEPGEHPDHVVVVKVPPVGDGSKAMDEYTSIEIFMGKNTIVLHNT 420
 QY 420 CEDSLAAPITILDVLAELSTRIOFKAENEGKFSHPVATILSYLTAKAPLVPPGTPV 479
 DB 421 CEDSLAAPITILDVLAELSTRIOFKAENEGKFSHPVATILSYLTAKAPLVPPGTPV 480
 QY 480 NALSKORAMLENTIRACVGLAPENNMLEYK 510
 DB 481 NALSKORAMLENTIRACVGLAPENNMLEYK 510
 RESULT 8
 ID INOI MSCR STANDARD; PRT: 512 AA.
 AC Q40271;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (Mt-1-P synthase) (1PS).
 DE Mesembryanthemum crystallinum (Common ice plant).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Aizoaceae; Mesembryanthemum.
 OC NCBI_TaxID=3544;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=96208959; PubMed=8624516;
 RA Ishitani M., Majumder A.L., Bornhouser A., Michalowski C.B.,
 RA Jensen R.G., Bohmelt H.J.;
 RT "Coordinate transcriptional induction of myo-inositol metabolism during environmental stress."
 RL Plant J. 9:537-548(1996).
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (by similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U32511; AAB03687.1; --
 DR PIR; T12438; T12438.
 DR HSSP; P11986; 1PIK.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 DR SEQUENCE 512 AA; 5675 MW; DF4B7109F47E8516 CRC64;
 Query Match 91.0%; Score 2395; DB 1; Length 512;
 Best Local Similarity 89.5%; Pred. No. 2.8e-149;
 Matches 458; Conservative 27; Mismatches 25; Indels 2; Gaps 1;
 QY 1 MFIEHFVSPVVKYETETETIOSVNYETTELVEHNRN--GTQWIVKPKSVNYQFKTTHV 58
 DB 1 MFIEHFVSPVVKYETETETIOSVNYETTELVEHNRKADAGVQWIVKPKSVNYQFKTTHV 60
 QY 59 VPKLGMLVKGNGNSTLTGVIANREDISWATKXIQOANFGSLTQASAIRVSGFOE 118
 DB 61 VPKLGMLVKGNGNSTLTGVIANREGISWATKXIQOANFGSLTQASAIRVSGFNG 120
 QY 119 EIIYAPFKSLPWNPPDDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESHVPLP 178
 DB 121 EIIYAPFKSLPWNPPDDIVFGGMDISNNMLADAMARAKVFDIDLOKOLRPYMESHVPLP 180
 QY 179 GIYPDFIAANOEBRANNVIKGTQOEVOQOIIKDIKAFKEATKVDKVVVLTANTERYSN 238
 DB 181 GIYPDFIAANOEBRANNVIKGTQOEVOQOIIKDIKAFKEATKVDKVVVLTANTERYSN 240
 QY 239 LVVGLNDTMENLLAVDRNEAISESTLYAIACWENVPFINGSPONTFVPGILDALAIR 298
 DB 241 VVGANDTMENLLAVDRNEAISESTLYAIACWENVPFINGSPONTFVPGILDALAIR 300
 QY 299 NTLIGGDDPKSGGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDSAPOTFRSKEIS 358
 DB 301 NTLIGGDDPKSGGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDSAPOTFRSKEIS 360
 QY 359 KSNVVDVNVSNALILYEPGEHPDHVVVVKVPPVGDGSKAMDEYTSIEIFMGKNTIVLHNT 418
 DB 361 KSNVVDVNVSNALILYEPGEHPDHVVVVKVPPVGDGSKAMDEYTSIEIFMGKNTIVLHNT 420
 QY 419 TCEDSLAAPITILDVLAELSTRIOFKAENEGKFSHPVATILSYLTAKAPLVPPGTPV 478
 DB 421 TCEDSLAAPITILDVLAELSTRIOFKAENEGKFSHPVATILSYLTAKAPLVPPGTPV 480
 QY 479 VNALSKORAMLENTIRACVGLAPENNMLEYK 510
 DB 481 VNALSKORAMLENTIRACVGLAPENNMLEYK 512
 RESULT 9

ID	INNO_2	ARATH	STANDARD	PRT	510 AA.
AC	Q38652	OS1E2			
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	25-OCT-2004	(Rel. 45, Last annotation update)			
DE	inositol-3-phosphate synthase isozyme 2 (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase 2) (M1-P synthase 2) (IFS 2).				
GN	OrderedLocNames=At2g22240				
OS	Arabidopsis thaliana (mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Columbia;				
RA	Johnson M.D., Burk D.H.;				
RT	Isozyme of 1L-myo-inositol-1-phosphate synthase from Arabidopsis."				
RL	(ex) Plant Gene Register PGR95-067.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Columbia;				
RA	Lin X., Kaul S., Rounsley S.D., Shea T.P., Bente M.-I., Town C.D.,				
RA	Fuili C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,				
RA	Buell C.R., Ketchum K.A., Lee J.D., Ronning C.M., Koo H.L.,				
RA	Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,				
RA	Tallon L.J., Gill J.E., Adams M.D., Carreira A.J., Creasy T.H.,				
RA	Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D.,				
RA	Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,				
RA	Venter J.C.;				
RT	"Sequence and analysis of chromosome 2 of the plant Arabidopsis				
RT	thaliana."				
RL	Nature 402:761-768(1999).				
CC	-1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-				
CC	phosphate.				
CC	-1- COPACTOR: NAD.				
CC	-1- PATHWAY: inositol biosynthesis.				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.				
CC	-1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase				
CC	family.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on leg-				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; U30250; AAC49172.1; -.				
DR	EMBL; AC007168; AAD23618.1; -.				
DR	PIR; D84610; D84610.				
DR	HSSP; P11986; 1PIJ.				
DR	InterPro; IPR002587; Inos-1-P synth.				
DR	Pfam; PF01658; Inos-1-P synth.1.				
KM	inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.				
FT	CONFLICT 135 135 E -> D (in Ref. 1).				
FT	CONFLICT 287 287 F -> L (in Ref. 1).				
FT	CONFLICT 298 298 K -> N (in Ref. 1).				
FT	CONFLICT 318 318 L -> W (in Ref. 1).				
FT	CONFLICT 477 477 P -> A (in Ref. 1).				
FT	CONFLICT 487 487 A -> P (in Ref. 1).				
SO	SEQUENCE 510 AA; 56337 MW; BDD58C166E289C85 CRC64;				
QY	Query Match 90.6%; Score 2384; DB 1; Length 510;				
	Best Local Similarity 88.2%; Pred. No. 1.5e-148;				
	Matches 450; Conservative 37; Mismatches 23; Indels 0; Gaps 0;				
DB	1 MFIEFKESNNVKTETETISVYNYETTELVHNRNCTQYIYKPKSVATQFKNTHP 60				
	1 MFISFKESNNVKTENEINSYDYETTELVEHNRNCTQYQVNVKPKVKDFKIDTFP 60				

Oy		61	KLGVLMLVGMGNGNSGLTGTGVIANRREDISVATKDDKIQQANYPFSLTQOASIRVSGEGEE	120
Dd		61	KLGVLMLVGMGNGNSGLTGTGVIANKEEISVATKDDKIQQANYPFSLTQOASSIRVSGYNGEE	120
Oy		121	IYAPEFKSLIPMVNPDDIVFGWMDISNNMNLADAMARAKVFEDIDLOKOLRPYMESMWPLPGI	180
Dd		121	IYAPEFKSLIPMVNPDEDFVGWMDISDNMLADAMARAVLIDIDLOKOLRPYEMENMIPLPGI	180
Oy		181	YDPPFIANAQBERRANNIKTKTOEQVOQLIKDKAPFEARKVDKKVVLMWTANTERYSNLY	240
Dd		181	YDPPFIANAGSRANSVIKTCKKEQVDHIIKDMREFEKVKVLLVMWTANTERSNNVI	240
Oy		241	VGLNDTMENILAAVDRREAEISPSSTIYAIAACMENVPFINGSPONFVPEGLIDLAIRART	300
Dd		241	VGLNDTMENILAAVEKDSEISPSSTIYAIACVLEGIFINGSQNFTVPEGLIELAISKNC	300
Oy		301	LIGGDDEFKSQTOKMKSVLVDFVLGAGIKPTSIYSVNYHLGNNDGMNISAPOTFRSKETSKS	360
Dd		301	LIGGDDEFKSQTOKMKSVLVDFVLGAGIKPTSIYSVNYHLGNNDGMNISAPOTFRSKETSKS	360
Oy		361	NVVDMDMNNAAILYPEBEPHDHVUVKVVPYVGDSDKRAMEYSELFMGSKNTIYLHNTC	420
Dd		361	NVVDMDMAASNGILFEPBEPDHVVUKVVPYVADSRRAMEYSELFMGRNTIYLHNTC	420
Oy		421	EDSLIAPRIILDVILAELSTRIOFKANEKGKHSFFHPVATIILSYLTKAFLVPPGTPPVNV	480
Dd		421	EDSLIAPRIILDVILAELSTRIOFKANEKGKHSFFHPVATIILSYLTKAFLVPPGTPPVNV	480
Oy		481	ALSKORAMLENINPACYGLAPENNMLEYK	510
Dd		481	ALSKORAMLENIILACVGLAPENNMTMEYK	510
<hr/>				
RESULT 10				
ID	INOI_BRANA	STANDARD;	PRT;	510 AA.
AC	O96348;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Inositol-3-phosphate synthase (EC 5.5.1.4) (myo-inositol-1-phosphate synthase) (M-1-P synthase) (IPS).			
OS	Brassica napus (Rape).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoids II; Brassicales; Brassicaceae; Brassica.			
OX	NCBI_TaxId=3708;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Huessein A., Bougeois J., Polvi S., Tsang E., Keller W.A., Georges F.;			
RT	"Cloning of a full length cDNA encoding myo-inositol 1-phosphate synthase from Brassica napus.";			
RL	Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.			
CC	-1 CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-phosphate.			
CC	-1 COFACTOR: NAD (By similarity).			
CC	-1 PATHWAY: Inositol biosynthesis.			
CC	-1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1 SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.			
<hr/>				
----- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announce/ or send an email to license@ebi.ac.uk). -----				
DR	EMBL; U66307; AAB06756.2; --			
DR	PIR; T08436; T08436.			
DR	HSSP; P11986; IPIJ.			

DR InterPro: IPR002587; Inos-1-P synth.
DR Pfam: PF01658; Inos-1-P synth. 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56377 MW; A40EB658DB80739 CRC64;

Query Match 90.3%; Score 2376; DB 1; Length 510;
Best Local Similarity 88.6%; Pred. No. 5e-148;
Matches 452; Conservative 29; Mismatches 29; Indels 0; Gaps 0;

QY 1 MFENFKYSPNVKYTEIETISQSVNYETTELHNRNGTYQVIKPKSVNTQFKNTNHP 60
DB 1 MFIESFKYSPNVKYTEIETISQSVNYETTELHNRNGTYQVIKPKSVNTQFKNTNHP 60
QY 61 KLGVMLVGMGNGNSTLTAGVIARREDISWATKDKIQOANFGSLTQASAIRVSGFOEE 120
DB 61 KLGVMLVGMGNGNSTLTAGVIARREGISWATKDKIQOANFGSLTQASAIRVSGFNEE 120
QY 121 IYAPFKSLPVMNPDPIYFGGMDISNMNLADAMARAKYFDIDLOKQLRPYMESVPLPGI 180
DB 121 MYAPFKSLPVMNPDPIYFGGMDISNMNLADAMARAKYFDIDLOKQLRPYMENIYPLPGI 180
QY 181 YDPDFIAANQSEERANNVIKGTKEQVOQIIKDIKAFKATKVDKVVMTANTERYSNLV 240
DB 181 YDPDFIAANQSEERANNVIKGTKEQVOQIIKDMREPKENKVDKVVMTANTERYSNVI 240
QY 241 VGLNDTMENLLAANDRNEAISPSTLYAIACMVNPFPINSPOPTFVPGILDLAIARNT 300
DB 241 VGLNDTMENLLANSVDRESEISPSSTLYAIACVLEGIPIFINSPOPTFVPGILDLAIKNNV 300
QY 301 LIGGDDFKSGGOTKMSVYVDFLVAGAGIKPSTISVYNHLGNDGNMLSAPOTFRSKEISKS 360
DB 301 LIGGDDFKSGGOTKMSVYVDFLVAGAGIKPSTISVYNHLGNDGNMLSAPOTFRSKEISKS 360
QY 361 NVYDDMNSNALIYEPGEHPHVVVIVKYVPYVGSKRAMEYTSIEFMGKNTVLAHNTC 420
DB 361 NVYDDMNSNGLIYEPGEHPHVVVIVKYVPYVADSKRAMDEYTSIEFMGKNTVLAHNTC 420
QY 421 EDLSLLAAPIIIDLVLAELSTRIOFKANEKGKFSFHHVATILSVLTQAPLPVPGTTPVN 480
DB 421 EDLSLLAAPIIIDLVLAELSTRIOFKANEKGKFSFHHVATILSVLTQAPLPVPGTTPVN 480
QY 481 ALSKORAMLENIMRACVGLAEPNNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAEPNNMILEYK 510

RESULT 11
INOS_1P synth. STANDARD; PRT; 510 AA.
AC Q9LX12; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Probable inositol-3-phosphate synthase isoform 3 (BC 5.5.1.4) (Myo-
inositol-1-phosphate synthase 3) (MT-1-P synthase 3) (IFS 3).
GN OrderedLocustNames=AL5910170; ORFNames=T31P16_160;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsida.
OC NCBI_TaxID=3702;
RA NCBI_TaxID=3702;
RA SEQUENCE FROM N.A.
RA STRAIN=cv. Columbia;
RA MEDLINE=21016721; PubMed=1130714; DOI=10.1038/35048507;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu B.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Katsumoto K., Matsuno A., Muraki A., Nakayama S.,
RA Nakatani N., Naito K., Okumura S., Shitipo S., Takeuchi C., Wada T.,
RA Wakanebe A., Yamada M., Sato S., de la Bastide M.,
RA Huang B., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stonerking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,

RA Belter B., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Molyneux E., Ozeraky P., Riley A., Stromstedt C.,
RA Wagner-McPherson C., Wollam A., Yoakum M., Bell W., Dedha N.,
RA Parnell L., Shah R., Rodriguez M., Hoan See L., Vail D., Baker J.,
RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,
RA Mariensen R., McCombe W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wandut R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Eutlaer K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grynoprez B., Zimmermann W.,
RA Ramberger U., Medler H., Balke K., Medler E., Peters S.,
RA van Steveren M., Dirkse W., Moeljan P., Klein Lankhorst R.,
RA Weizenecker T., Bothe G., Rose M., Hauf J., Bernaisser S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bente O., Lemcke K., Kolesov G., Mayer K.F.X., Rüd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
thaliana".
RL Nature 408:823-826(2000).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL356332; CAB92058.1; -.
DR PIR: T50021; T50021.
DR HSSD: P1986; 1PIH.
DR InterPro: IPR002587; Inos-1-P synth.
DR Pfam: PF01658; Inos-1-P synth. 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56417 MW; 5CB8108082152473 CRC64;

Query Match 90.3%; Score 2376; DB 1; Length 510;
Best Local Similarity 87.8%; Pred. No. 5e-148;
Matches 448; Conservative 40; Mismatches 22; Indels 0; Gaps 0;

QY 1 MFENFKYSPNVKYTEIETISQSVNYETTELHNRNGTYQVIKPKSVNTQFKNTNHP 60
DB 1 MFIESFKYSPNVKYTEIETISQSVNYETTELHNRNGTYQVIKPKSVNTQFKNTNHP 60
QY 61 KLGVMLVGMGNGNSTLTAGVIARREDISWATKDKIQOANFGSLTQASAIRVSGFOEE 120
DB 61 KLGVMLVGMGNGNSTLTAGVIARREGISWATKDKIQOANFGSLTQASAIRVSGFNEE 120
QY 121 IYAPFKSLPVMNPDPIYFGGMDISNMNLADAMARAKYFDIDLOKQLRPYMESVPLPGI 180
DB 121 IYAPFKSLPVMNPDPIYFGGMDISNMNLADAMARAKYFDIDLOKQLRPYMENIYPLPGI 180
QY 181 YDPDFIAANQSEERANNVIKGTKEQVOQIIKDIKAFKATKVDKVVMTANTERYSNLV 240
DB 181 YDPDFIAANQSEERANNVIKGTKEQVOQIIKDMREPKENKVDKVVMTANTERYSNVI 240
QY 241 VGLNDTMENLLAANDRNEAISPSTLYAIACMVNPFPINSPOPTFVPGILDLAIARNT 300
DB 241 VGLNDTMENLLANSVDRESEISPSSTLYAIACVLEGIPIFINSPOPTFVPGILDLAIKNNV 300
QY 301 LIGGDDFKSGGOTKMSVYVDFLVAGAGIKPSTISVYNHLGNDGNMLSAPOTFRSKEISKS 360
DB 301 LIGGDDFKSGGOTKMSVYVDFLVAGAGIKPSTISVYNHLGNDGNMLSAPOTFRSKEISKS 360
QY 361 NVYDDMNSNALIYEPGEHPHVVVIVKYVPYVGSKRAMEYTSIEFMGKNTVLAHNTC 420
DB 361 NVYDDMNSNGLIYEPGEHPHVVVIVKYVPYVADSKRAMDEYTSIEFMGKNTVLAHNTC 420

Dh 361 NVDDMVGSGNCLYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVMNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Dh 421 EDSLAAPIIIDLVLLAELSTRIOFKMSNEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Dh 481 ALSKORAMLENIVLRACVGLAPENNMLEYK 510

RESULT 12

Q9AR12 PRELIMINARY; PRT; 509 AA.
ID Q9AR12
AC Q9AR12
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2003 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Myo-inositol 1-phosphate synthase.
OS Avicennia marina (Grey mangrove).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Acanthaceae; Acanthaceae; Acanthaceae; Acanthaceae; Acanthaceae;
NCBI_TaxID=82927;
RN [1]
RP SEQUENCE FROM N.A.
RA Jithesh M.N., Parani M., Parida A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028259; AAK21969.1; -.
DR HSSP; P11986; 1PK.
DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR02587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth. 1.
SQ SEQUENCE 509 AA; 55978 MW; 23C8D354BAF3BD0F CRC64;

Query Match 90.2%; Score 2373.5; DB 2; Length 509;
Best Local Similarity 89.6%; Pred. No. 7.3e-148;
Matches 457; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

Qy 1 MFIEFKVESPNVKTETETISQVNVYETTELVEHNRNGTQWIVKPSVNTQFKTNTHVP 60
Dh 1 MFIEFKVESPNVKTETETISQVNVYETTELVEHNRNGTQWIVKPSVNTQFKTNTHVP 60
Qy 1 KLGWLVGMGNGNSTLTGVIANREDISWATKDIQOANYFGLTQASAIRVGSFOGEE 120
Dh 1 KLGWLVGMGNGNSTLTGVIANREDISWATKDIQOANYFGLTQASAIRVGSFOGEE 120
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDIQOANYFGLTQASAIRVGSFOGEE 120
Dh 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDIQOANYFGLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPVMNPDIDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPVYESVPLPGI 180
Dh 121 IYAPFKSLPVMNPDIDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPVYESVPLPGI 180
Qy 181 YDPDFIAANOGEERANNVTKGTKEQVOQI IKDIAFKFKAATKYVDKVVVLTANTERYSNLV 240
Dh 181 YDPDFIAANOGEERANNVTKGTKEQVOQI IKDIAFKFKAATKYVDKVVVLTANTERYSNLV 240
Qy 181 YYPDFIAANOGSRANNVTKGTKEQVOQI IKDIAFKFKAATKYVDKVVVLTANTERYSNLV 239
Dh 181 YYPDFIAANOGSRANNVTKGTKEQVOQI IKDIAFKFKAATKYVDKVVVLTANTERYSNLV 239
Qy 241 VGLNDTMENLLAAVDNRNAAETISPTLYAIACVMEVNPFIINGSPONTFVPGILDLAIARNT 300
Dh 241 VGLNDTMENLLAAVDNRNAAETISPTLYAIACVMEVNPFIINGSPONTFVPGILDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Dh 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 300 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPQTFRSKEISKS 359
Dh 300 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPQTFRSKEISKS 359
Qy 361 NVDDMVSNSAILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVLANTC 420
Dh 361 NVDDMVSNSAILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVLANTC 420
Qy 360 NVDDMVSNSAILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVLANTC 419
Dh 360 NVDDMVSNSAILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVLANTC 419
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Dh 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Qy 420 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 479
Dh 420 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 479

Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Dh 480 ALSKORAMLENIVLRACVGLAPENNMLEYK 509

RESULT 13

Q7XUCO PRELIMINARY; PRT; 510 AA.
ID Q7XUCO
AC Q7XUCO
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Myo-inositol 1-phosphate synthase INOI.
OS Xerophyta viscosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Velloziaceae; Xerophyta.
NCBI_TaxID=90708;
RN [1]
RP SEQUENCE FROM N.A.
RA Majee M., Majumder A.N.L., Munder S.G.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY323824; AAP8551.1; -.
DR HSSP; P11986; 1UKI.
DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR02587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth. 1.
SQ SEQUENCE 510 AA; 56234 MW; 5F92212651115A2A CRC64;

Query Match 89.7%; Score 2361; DB 2; Length 510;
Best Local Similarity 88.6%; Pred. No. 4.8e-147;
Matches 452; Conservative 28; Mismatches 30; Indels 0; Gaps 0;

Qy 1 MFIEFKVESPNVKTETETISQVNVYETTELVEHNRNGTQWIVKPSVNTQFKTNTHVP 60
Dh 1 MFIEFKVESPNVKTETETISQVNVYETTELVEHNRNGTQWIVKPSVNTQFKTNTHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDIQOANYFGLTQASAIRVGSFOGEE 120
Dh 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDIQOANYFGLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPVMNPDIDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPVYESVPLPGI 180
Dh 121 IYAPFKSLPVMNPDIDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPVYESVPLPGI 180
Qy 181 YDPDFIAANOGEERANNVTKGTKEQVOQI IKDIAFKFKAATKYVDKVVVLTANTERYSNLV 240
Dh 181 YDPDFIAANOGEERANNVTKGTKEQVOQI IKDIAFKFKAATKYVDKVVVLTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDNRNAAETISPTLYAIACVMEVNPFIINGSPONTFVPGILDLAIARNT 300
Dh 241 VGLNDTMENLLAAVDNRNAAETISPTLYAIACVMEVNPFIINGSPONTFVPGILDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Dh 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Dh 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVSNSAILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVLANTC 420
Dh 361 NVDDMVSNSAILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVLANTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Dh 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Dh 481 ALSKORAMLENIVLRACVGLAPENNMLEYK 510

RESULT 14

INOI_SPIPO

```

ID INOI SPIPO STANDARD; PRT; 510 AA.
AC P42803;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (MI-1-P synthase) (IPS).
GN Name=TurI;
OS Spirodela polyrrhiza (Giant duckweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Lemnoidae;
OC Spirodela.
OX NCBI_TaxID=29656;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94035182; PubMed=8220483;
RA Smart C.C., Fleming A.J.;
RT "A plant gene with homology to D-myo-inositol-3-phosphate synthase is
RT rapidly and spatially up-regulated during an abscisic-acid-induced
RT morphogenic response in Spirodela polyrrhiza.";
RL Plant J. 4:279-293(1993).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- INDUCTION: By abscisic acid (ABA).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z11693; CAAT7751.1; -
DR PIR; S60302; S60302.
DR HSSP; P11986; P1H.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56385 MW; 2D56D3666C5E03C CRC64;

Query Match 89.6%; Score 2358; DB 1; Length 510;
Best Local Similarity 87.8%; Pred. No. 7,6e-147;
Matches 448; Conservative 37; Mismatches 25; Indels 0; Gaps 0;

OY 1 MEIENFKYESPNVKTETETIQSIVNYETTELVHENRNGTYOMIVPKSVNYQFKTNHP 60
DB 1 MEIEKFRVESPNVKKYGDGIESVSYETTELVEHVRNGSYQWVVRPKSVQYQFKTDTRVP 60
OY 61 KLGVNLVGMGGNGSTLTGVIANREDISWATKDIQOANYGSGITQASAPRVSGOBE 120
DB 61 KLGVNLVGMGGNGSTLTGVIANREGISWATKEKYQOANFGSLTQSSSIRVSGFNGBE 120
OY 121 IYAPFKSLPVMVPPDIVFGGWDISNNMLADAMARAKYFDIDLOQLRPYMESWPLPGI 180
DB 121 IYAPFKSLPVMVPPDIVFGGWDISNNMLADAMARAKYFDIDLOQLRPYMESWPLPGI 180
OY 121 IYAPFKSLPVMVPPDIVFGGWDISNNMLADAMARAKYFDIDLOQLRPYMESWPLPGI 180
DB 121 IYAPFKSLPVMVPPDIVFGGWDISNNMLADAMARAKYFDIDLOQLRPYMESWPLPGI 180
OY 181 YDPDFIAANQSERANNAVITKTKQBOVOQIIDIKAPKEATKYDKVVTMTANTERYSNLV 240
DB 181 YDPDFIAANQSERANNAVITKTKQBOVOQIIDIKAPKEATKYDKVVTMTANTERYSNLV 240
OY 181 YNPDPFIAANQSRANNAVITKTKQBOVOQIIDIKAPKEATKYDKVVTMTANTERYSNLV 240
DB 181 YNPDPFIAANQSRANNAVITKTKQBOVOQIIDIKAPKEATKYDKVVTMTANTERYSNLV 240
OY 241 VGLNDTMEENLAAVDRNEAISPSTLYAIACVMEVPPFINSPOPTFVGLIDLAIABNT 300
DB 241 VGLNDTMEENLAAVDRNEAISPSTLYAIACVMEVPPFINSPOPTFVGLIDLAIABNT 300
OY 241 VGLNDTMEENLAAVDRNEAISPSTLYAIACVMEVPPFINSPOPTFVGLIDLAIABNT 300
DB 241 VGLNDTMEENLAAVDRNEAISPSTLYAIACVMEVPPFINSPOPTFVGLIDLAIABNT 300
OY 301 LIGGDDFKSGGOTKMSVLVDPLVGAIGIKPTSIIVSYNHLGNDGNMLSAAPOTFRSKEISKS 360
DB 301 LIGGDDFKSGGOTKMSVLVDPLVGAIGIKPTSIIVSYNHLGNDGNMLSAAPOTFRSKEISKS 360

```

```

OY 361 NVVDDMNSNALIVRGSHPHVIVIKYVPYVGSKRAMDEYTSRIFMGKRTYILAHNTC 420
DB 361 NVVDDMNSNGLIVRGSHPHVIVIKYVPYVGSKRAMDEYTSRIFMGKSTIILAHNTC 420
OY 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAAPLVPPGTPPVN 480
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTAAPLVPPGTPPVN 480
OY 481 ALSKRAMLENTIRACVGLAPENNMLEYK 510
DB 481 ALSKRAMLENTIRACVGLAPENNMLEYK 510

RESULT 15
ID INOI MA1ZE STANDARD; PRT; 510 AA.
AC 09PFK7; O65196;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (MI-1-P synthase) (IPS).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Early ACR; TISSUE=leaf;
RA Larson S.R., Raboy V.;
RT "Linkage mapping maize and barley myo-inositol 1-phosphate synthase
RT genes.";
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX Shukla S., Vantsoi T.T.;
RT "Genomic sequence of maize myo-inositol 1-phosphate synthase gene.";
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF056326; AAC15756.1; -
DR EMBL; AF323175; AAG40328.1; -
DR PIR; T01647; T01647.
DR HSSP; P11986; P1H.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
KW CONFLICT 18 M -> T (in Ref. 2).
FT CONFLICT 351 A -> T (in Ref. 2).
SQ SEQUENCE 510 AA; 56245 MW; DAB59EECF391CB6D CRC64;

Query Match 89.5%; Score 2354; DB 1; Length 510;
Best Local Similarity 88.2%; Pred. No. 1.4e-146;
Matches 450; Conservative 27; Mismatches 33; Indels 0; Gaps 0;

OY 1 MEIENFKYESPNVKTETETIQSIVNYETTELVHENRNGTYOMIVPKSVNYQFKTNHP 60
DB 1 MEIENFKYESPNVKTETETIQSIVNYETTELVHENRNGTYOMIVPKSVNYQFKTNHP 60

```

```
Db      1 MFIESFRVSPHVRXGYPMIESEYRYDTTELVEHGKGASRWVVRPKSVKYNFRTRTAVP 60
Qy      61 KLGWLVGMGNGNGSTLTGVTIANREDISWATKDKIOOANYFGSLTQASAIRVGSFOGEE 120
Db      61 KLGWLVGMGNGNGSTLTGVTIANREGISWATKDKVOOANYGSLTQASTIRVGSYNSEE 120
Qy      121 IYAPFKSLLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQOLRPYMESWVPLPGI 180
Db      121 IYAPFKSLLPMVNPDDIVFGGWDISNNMLADSMTRAKVLIDLQOLRPYMESWVPLPGI 180
Qy      181 YDPPEIANOEERANNVVKGTQOEVOQI IKDIAKFEATKVDKVVLTANTERYSNLV 240
Db      181 YDPPEIAANQSRANSVIKGTKEQVEQI IKDIREFEKENKVDKIVLMTANTERYSNVC 240
Qy      241 VGLNDTMENLAAVDRNEAEISPTLYAIACMENVPFINGSPONTFVPGIIDLAIANT 300
Db      241 AGLNDTMENLAAVDKNEAEVSPSTLYAIACMEGVPPINGSPONTFVPGIIDLAIKNNC 300
Qy      301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSI VSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db      301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSI VSYNHLGNNDGMNLSAPQAFRSKEISKS 360
Qy      361 NVUDDMVNSNAI LYERGEHPDHVVVVKYVPYVGDSKRAMDEYTSRI FMGKQTI VLAHNTC 420
Db      361 NVUDDMVNSNAI LYERGEHPDHVVVVKYVPYVGDSKRAMDEYTSRI FMGKQTI VLAHNTC 420
Qy      421 EDSLILAPIIIDLVLAELESTRIQFKANEKGKFSFHPVATILSYLTAKAPLVPFGTPVYN 480
Db      421 EDSLILAPIIIDLVLAELESTRIQFKABEGDEKFSFHPVATILSYLTAKAPLVPFGTPVYN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALAKORAMLENIMRACVGLAPENNMILEYK 510
```

Search completed: June 7, 2005, 16:46:41
Job time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:16:29 ; Search time 24.5 Seconds
(without alignments)
2002.860 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MEIEFKVESPVVKYTEIEI.....NIMRACVGLAPENNILEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2395	91.0	512	2 T12438	inositol-3-phospha
2	2384	90.6	510	2 D84610	probable myo-inos
3	2376	90.3	510	2 T50021	inositol-3-phospha
4	2371	90.1	509	2 T08436	inositol-3-phospha
5	2358	89.6	510	2 S60302	inositol-3-phospha
6	2354	89.5	510	2 T01647	inositol-3-phospha
7	2349.5	89.3	511	2 T05017	inositol-3-phospha
8	2338	88.9	510	2 T04399	inositol-3-phospha
9	2304.5	87.6	507	2 T10964	inositol-3-phospha
10	2304.5	87.6	507	2 S52648	inositol-3-phospha
11	1393.5	53.0	525	2 T18569	inositol-3-phospha
12	1389.5	52.8	520	2 S45452	inositol-3-phospha
13	1389.5	52.8	555	2 A30902	inositol-3-phospha
14	1312.5	49.9	430	2 T46317	inositol-3-phospha
15	291	11.1	388	2 T36191	probable myo-inos
16	284	10.8	392	2 A69474	myo-inositol-1-pho
17	233	8.9	417	2 T34930	probable secreted
18	215	8.2	382	2 H72255	myo-inositol-1-pho
19	214	8.1	425	2 E71039	hypothetical prote
20	197	7.5	386	2 B75175	myo-inositol-1-pho
21	193	7.3	364	2 G70451	conserved hypothet
22	161	6.1	360	2 B90239	conserved hypothet
23	156	5.9	392	2 F72632	probable myo-inos
24	143.5	5.5	368	2 A69014	conserved hypothet
25	142	5.4	369	2 S72835	hypothetical prote
26	132	5.0	360	2 T36586	hypothetical prote
27	127.5	4.8	367	2 F70912	hypothetical prote
28	124.5	4.7	367	2 G95871	conserved hypothet
29	120.5	4.6	1993	2 AF1450	probable peptidogl

30	118.5	4.5	853	2 G90559	tree-like protein
31	118	4.5	731	2 D95235	pecticillin-binding
32	118	4.5	802	2 C9326	hypothetical prote
33	117.5	4.5	2481	2 A43908	fibronectin - Afri
34	116	4.4	1127	2 T28317	ORF MSV156 hypothe
35	115.5	4.4	452	2 S77155	UDP-N-acetylmuramo
36	115.5	4.4	873	2 T16282	hypothetical prote
37	114.5	4.4	344	2 B90176	conserved hypothet
38	114	4.3	731	2 E98099	peptidoglycan glyco
39	113.5	4.3	1628	2 E90538	hypothetical prote
40	113.5	4.3	1792	2 T20363	hypothetical prote
41	113	4.3	867	2 H90524	preprotein transio
42	112.5	4.3	900	2 S46101	hypothetical prote
43	112	4.3	2490	1 A54971	protein-tyrosine-p
44	111.5	4.2	422	2 C89953	isocitrate dehydrog
45	111.5	4.2	867	2 B86815	ClpB protein (limpo

ALIGNMENTS

RESULT 1

T12438 inositol-3-phosphate synthase (EC 5.5.1.4) - common ice plant

C:Species: Mesembryanthemum crystallinum (common ice plant)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004

C:Accession: T12438

R:Rishitani, M.; Majumder, A.L.; Bornhouser, A.; Michalowski, C.B.; Jensen, R.G.; Bohner, R.

Plant J. 9, 537-548, 1996

A:Title: Coordinate transcriptional induction of myo-inositol metabolism during environme

A:Reference number: Z17518; MUID:96208959; PMID:8624516

A:Accession: T12438

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-512 <ISH>

A:Cross-references: UNIPROT:Q40271; EMBL:U32511; NID:g975887; PIDN:AAB03687.1; PID:g97588

C:Superfamily: myo-inositol-1-phosphate synthase

C:Keywords: intramolecular lyase; isomerase

Query Match 91.0%; Score 2395; DB 2; Length 512;

Best Local Similarity 89.5%; Pred. No. 3,1e-153;

Matches 458; Conservative 27; Mismatches 25; Indels 2; Gaps 1;

QY	1	MEIEFKVESPVVKYTEIEIQSVYVETTELVEHNRN--GTYQWVYKSVYVQKRTNH	58
DB	1	MEIEFKVESPVVKYTEIEISVYVDTTELVEHNRKDAGGQWVYKRTVQYHFKTDTR	60
QY	59	VFKLGVMLVWGKNGNSTLTGTVIANREDISWATKDKIQANVFGSLTQASAIRVGSFQG	118
DB	61	VFKLGVMLVWGKNGNSTLTGTVIANREGISWATKDKIQANVFGSLTQASAIRVGSFQG	120
QY	119	BEIVAPFKSLPMVNPDDIVFGMDISNMNLADAMARAKVPDIDLOQLRPYMSMVPPL	178
DB	121	BEIVAPFKSLPMVNPDDIVFGMDISNMNLADAMARAKVPDIDLOQLRPYMSMVPPL	180
QY	179	GIYDPDFIAANOEEAANVYIKTKQOYQIIKIKAKKATKDKVYVLTANTERISN	238
DB	181	GIYDPDFIAANOEEAANVYIKTKQOYQIIKIKAKKATKDKVYVLTANTERISN	240
QY	239	LVGAGNDMENLAAVDNBEAISPSTLYATACVWENVPFNGSPQNTFVPLDILATAR	298
DB	241	VVGAGNDMENLAAVDNBEAISPSTLYATACVWENVPFNGSPQNTFVPLDILATAR	300
QY	299	NLTIGDDFKSGQTKMSVLVDPLVAGIKPTSIYSVYHLAGNDGMLASQTFERSKIS	358
DB	301	NSLTIGDDFKSGQTKMSVLVDPLVAGIKPTSIYSVYHLAGNDGMLASQTFERSKIS	360
QY	359	KSNVVDVNVSNALIIYEPGEHPDHVVYIKYIPYGDSSRAADETSETFNGSKATTVLHN	418
DB	361	KSNVVDVNVSNALIIYEPGEHPDHVVYIKYIPYGDSSRAADETSETFNGSKATTVLHN	420
QY	419	TCEDSLAAPIIIDVLAASHSTRIQFAENEGKPHSPVATISYVTKAPLVPGTPIV	478

Db 421 TCEDSLAAPIILDVLLAEISTRIOQLKAEEDKFSFHPVATILSYLTKAPLVPGTPV 480
Qy 479 VNALSKORAMLENIIRACVGLAPENNMILEYK 510
Db 481 VNALSKORAMLENIIRACVGLAPENNMILEYK 512

RESULT 2
D84610
probable myo-inositol 1-phosphate synthase [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004
C/Accession: D84610
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon,
L.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventner, J.
Nature 402, 761-766, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: D84610
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-510 <STO>
A/Cross-references: UNIPROT:Q38862; GB:AE002093; NID:g4567202; PTDN:AA023618.1; GSPDB:GN
A/Genetics:
A:Gene: At2g22240
A:Map position: 2
C/Superfamily: Myo-inositol-1-phosphate synthase

Query Match 90.6%; Score 2384; DB 2; Length 510;
Best Local Similarity 88.2%; Pred. No. 1.7e-152;
Matches 450; Conservative 37; Mismatches 23; Indels 0; Gaps 0;

Qy 1 MFIEPFKESPNVXYTETETESVYVYETTELVEHNRGTQWIVPKSVNYQFKTNTHVP 60
Db 1 MFIEPFKESPNVXYTETETESVYVYETTELVEHNRGTQWIVPKSVNYQFKTDIVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIQOANFGSLTQASAIRVGSFOGEE 120
Db 61 KLGWLVGMGNGNSTLTGVIANKEGISWATKDVQOANFGSLTQASAIRVGSYNSEE 120
Qy 121 IYAPFKSLPWNVPDDIVFGGWDISNNMLADAMARAKYFDIDLOKQLRPYVESWVPLPGI 180
Db 121 IYAPFKSLPWNVPDDIVFGGWDISNNMLADAMARAVLDIDLOKQLRPYEMNMLPLPGI 180
Qy 181 YDPDFIAANOEERANNVKGTKQEQVQIIKDIKAFKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANQSGRANVKGTKKEQVDHIKDMREFEKKNKVDKVVLMTANTERYSNVI 240
Qy 241 VGLNDTMENLLAADRNEAIESPTLYAIACVMENVPFINGSPONTFVPGILDLAIANT 300
Db 241 VGLNDTMENLLASVKESEISPTLYAIACVLEBIPFINGSPONTFVPGILELAISNC 300
Qy 301 LIIGDDPKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNMLAPQTFRSKEISKS 360
Db 301 LIIGDDPKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNMLAPQTFRSKEISKS 360
Qy 361 NVVDDMVNSNAILYEPGHPDHVVVVKYVPVYVGSKRAMDEYTSEIFMGKNTTILHNTC 420
Db 361 NVVDDMVNSNAILYEPGHPDHVVVVKYVPVYVGSKRAMDEYTSEIFMGKNTTILHNTC 420
Qy 421 EDSLAAPIILDVLLAEISTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTPV 480
Db 421 EDSLAAPIILDVLLAEISTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTPV 480
Qy 481 ALSKORAMLENIIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIIRACVGLAPENNMILEYK 510

RESULT 3
T50021
inositol-3-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis thaliana

N/Alternate names: protein T31P16.160
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C/Accession: T50021
R/Beyan, M.; Zimmermann, W.; Grueneisen, A.; Wandt, R.; Kalicki, J.; Woldmann, P.; Smi
submitted to the Protein Sequence Database, May 2000
A/Reference number: 225027
A/Accession: T50021
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-510 <BEV>
A/Cross-references: UNIPROT:Q9LX12; EMBL:AJ356332; GSPDB:GN00063; ATSP:T31P16.160
A/Experimental source: cultivar Columbia; BAC clone T31P16
C/Genetics:
A:Gene: ATSP:T31P16.160
A/Introns: 63/2; 86/3; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 90.3%; Score 2376; DB 2; Length 510;
Best Local Similarity 87.8%; Pred. No. 5.9e-152;
Matches 448; Conservative 40; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MFIEPFKESPNVXYTETETESVYVYETTELVEHNRGTQWIVPKSVNYQFKTNTHVP 60
Db 1 MFIEPFKESPNVXYTETETESVYVYETTELVEHNRGTQWIVPKSVNYQFKTDIVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIQOANFGSLTQASAIRVGSFOGEE 120
Db 61 KLGWLVGMGNGNSTLTGVIANREBISWATKDVQOANFGSLTQASAIRVGSFNGEE 120
Qy 121 IYAPFKSLPWNVPDDIVFGGWDISNNMLADAMARAKYFDIDLOKQLRPYVESWVPLPGI 180
Db 121 IYAPFKSLPWNVPDDIVFGGWDISNNMLADAMARAVLDIDLOKQLRPYEHVPLPGI 180
Qy 181 YDPDFIAANOEERANNVKGTKQEQVQIIKDIKAFKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANQSGRANVKGTKKEQVQIIKDIKAFKATKVDKVVLMTANTERYSNV 240
Qy 241 VGLNDTMENLLAADRNEAIESPTLYAIACVMENVPFINGSPONTFVPGILDLAIANT 300
Db 241 VGLNDTMENLLASLDKDEAIESPTLYAIACVLEBIPFINGSPONTFVPGILELAIRNC 300
Qy 301 LIIGDDPKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNMLAPQTFRSKEISKS 360
Db 301 LIIGDDPKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNMLAPQTFRSKEISKS 360
Qy 361 NVVDDMVNSNAILYEPGHPDHVVVVKYVPVYVGSKRAMDEYTSEIFMGKNTTILHNTC 420
Db 361 NVVDDMVNSNAILYEPGHPDHVVVVKYVPVYVGSKRAMDEYTSEIFMGKNTTILHNTC 420
Qy 421 EDSLAAPIILDVLLAEISTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTPV 480
Db 421 EDSLAAPIILDVLLAEISTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTPV 480
Qy 481 ALSKORAMLENIIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIIRACVGLAPENNMILEYK 510

RESULT 4
T08436
inositol-3-phosphate synthase (EC 5.5.1.4) [similarity] - rape
C/Species: Brassica napus (rape)
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: T08436
R/Husain, A.; Bourgeois, J.; Polvi, S.; Tsang, E.; Keller, W.A.; Georges, F.
submitted to the EMBL Data Library, August 1996
A/Reference number: 216418
A/Accession: T08436
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA

A:Residues: 1-509 <HUS>
A:Cross-references: UNIPROT:Q96348; EMBL:U66307; NID:g1513227; PID:g1513228
C:Function: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo
A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 90.1%; Score 2371; DB 2; Length 509;
Best Local Similarity 88.6%; Pred. No. 1.3e-151;
Matches 451; Conservative 29; Mismatches 29; Indels 0; Gaps 0;

QY 2 FIENKVESPNVKYETEYETELVHNRNGTYQWIVKPKSVNYQFKTNTVP 61
DB 1 FIESKVESPNVKYETENHSHYDYTEVHENVNGAYQWIVKPKVYDXTDTRVPK 60
QY 62 LGVNLVGMGNNGSTLTGVIANREDISWATKDIQOANYFGSLTQASIRVGSFGEEI 121
DB 61 LGVNLVGMGNNGSTLTGVIANKEGISWATKDIQOANYFGSLTQASIRVGSFGEEI 120
QY 122 YAPFKSLPMPNPDIVFGMDISNMNLADAMARAKVPIDLOKQLRPMESVPLPGI 181
DB 121 YAPFKSLPMPNPDIVFGMDISNMNLADAMARAKVLDIDLOKQLRPMENIVPLPGI 180
QY 182 DPDFIAANOEBRANNVIKGTKEQVOQIIKDIKAFKATKVDKVVVLTANTERTSNLV 241
DB 181 DPDFIAANOGBRANNVIKGTKEQVOQIIKDKREKREKVKVLTANTERTSNLV 240
QY 242 GLNDTMENLLAANDNEAEISPTLVAIACWENVPFINGSPOINTFVGLIDLAIAANT 301
DB 241 GLNDTMENLLANSVDDESEISPTLVAIACVLEGIPIFINGSPOINTFVGLIDLAIAKNVL 300
QY 302 IGGDFKSGQTKMKSVDPLVGAIGIKPTISVSNHLGNNDGMNLAPQTFRSKEISKS 361
DB 301 IGGDFKSGQTKMKSVDPLVGAIGIKPTISVSNHLGNNDGMNLAPQTFRSKEISKS 360
QY 362 VVDMDVMSNAIYEPGEHPDHVVVYKVPYVGSKRAMDEYTSIEFMGKNTIYLAHNTC 421
DB 361 VVDMDVMSNAIYEPGEHPDHVVVYKVPYVADSKRAMDEYTSIEFMGKNTIYLAHNTC 420
QY 422 DSLAAPRIIDLVLLAEISTRIQFAENEGKFSFHPVATILSYLTAKPLVPPTPVNA 481
DB 421 DSLAAPRIIDLVLLAEISTRIQFAENEGKFSFHPVATILSYLTAKPLVPPTPVNA 480
QY 482 LSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 LSKORAMLENIMRACVGLAPENNMILEYK 509

RESULT 5

S60302
inositol-3-phosphate synthase (EC 5.5.1.4) - Spirodela polyrrhiza
C:Species: Spirodela polyrrhiza
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60302
R:Smart, C.C.; Fleming, A.J.
Plant J. 4, 279-293, 1993
A:Title: A plant gene with homology to D-myo-inositol-3-phosphate synthase is rapidly an
A:Reference number: S60302; MUID:94035182; PMID:8220483
A:Accession: S60302
A:Molecule type: mRNA
A:Residues: 1-510 <SMA>
A:Cross-references: UNIPROT:P42803; EMBL:Z11693; NID:g396209; PIDN:CAA77751.1; PID:g5586
A:Gene: turtl
C:Genetics:
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 89.6%; Score 2358; DB 2; Length 510;
Best Local Similarity 87.8%; Pred. No. 9.5e-151;
Matches 448; Conservative 37; Mismatches 25; Indels 0; Gaps 0;

QY 1 MFIEFKVESPNVKYETEYETELVHNRNGTYQWIVKPKSVNYQFKTNTVP 60
DB 1 MFIEFKVESPNVKYETEYETELVHNRNGTYQWIVKPKSVNYQFKTNTVP 60

DB 1 MFIEFKVESPNVKYDGEIESYETTELVHEVRNGYQWVVKPKSVQYQFKTDTVP 60
QY 61 KLGVNLVGMGNNGSTLTGVIANREDISWATKDIQOANYFGSLTQASIRVGSFGEEI 120
DB 61 KLGVNLVGMGNNGSTLTGVIANREDISWATKDIQOANYFGSLTQASIRVGSFGEEI 120
QY 121 IYAPFKSLPMPNPDIVFGMDISNMNLADAMARAKVPIDLOKQLRPMESVPLPGI 180
DB 121 IYAPFKSLPMPNPDIVFGMDISNMNLADAMARAKVLDIDLOKQLRPMESVPLPGI 180
QY 241 YDPDFIAANOEBRANNVIKGTKEQVOQIIKDIKAFKATKVDKVVVLTANTERTSNLV 240
DB 241 YNPDFIAANOGBRANNVIKGTKEQVOQIIKDIKAFKATKVDKVVVLTANTERTSNLV 240
QY 242 GLNDTMENLLAANDNEAEISPTLVAIACWENVPFINGSPOINTFVGLIDLAIAANT 300
DB 241 GLNDTMENLLAANDNEAEISPTLVAIACWENVPFINGSPOINTFVGLIDLAIAANT 300
QY 301 IGGDFKSGQTKMKSVDPLVGAIGIKPTISVSNHLGNNDGMNLAPQTFRSKEISKS 360
DB 301 IGGDFKSGQTKMKSVDPLVGAIGIKPTISVSNHLGNNDGMNLAPQTFRSKEISKS 360
QY 361 NVDDMDVMSNAIYEPGEHPDHVVVYKVPYVGSKRAMDEYTSIEFMGKNTIYLAHNTC 420
DB 361 NVDDMDVMSNAIYEPGEHPDHVVVYKVPYVGSKRAMDEYTSIEFMGKNTIYLAHNTC 420
QY 421 EDLSLAAPRIIDLVLLAEISTRIQFAENEGKFSFHPVATILSYLTAKPLVPPTPVNA 480
DB 421 EDLSLAAPRIIDLVLLAEISTRIQFAENEGKFSFHPVATILSYLTAKPLVPPTPVNA 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 6

T01647
inositol-3-phosphate synthase (EC 5.5.1.4) - maize
C:Species: Zea mays (maize)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01647
R:Larson, S.R.; Raboy, V.
submitted to the EMBL Data Library, March 1998
A:Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A:Reference number: Z14366
A:Accession: T01647
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-510 <LAR>
A:Cross-references: UNIPROT:Q9P6K7; EMBL:AF056326; NID:g3108052; PIDN:AAIC15756.1; PID:g33
A:Experimental source: strain Early ACR; leaf
A:Genetics:
A:Gene: INO1
C:Function:
A:Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-ino
A:Pathway: inositol biosynthesis
A>Note: NAD cofactor
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 89.5%; Score 2354; DB 2; Length 510;
Best Local Similarity 88.2%; Pred. No. 1.8e-150;
Matches 450; Conservative 27; Mismatches 33; Indels 0; Gaps 0;

QY 1 MFIEFKVESPNVKYETEYETELVHNRNGTYQWIVKPKSVNYQFKTNTVP 60
DB 1 MFIEFKVESPNVKYETENHSHYDYTEVHENVNGAYQWIVKPKVYDXTDTRVP 60
QY 61 KLGVNLVGMGNNGSTLTGVIANREDISWATKDIQOANYFGSLTQASIRVGSFGEEI 120
DB 61 KLGVNLVGMGNNGSTLTGVIANREDISWATKDIQOANYFGSLTQASIRVGSFGEEI 120
QY 121 IYAPFKSLPMPNPDIVFGMDISNMNLADAMARAKVPIDLOKQLRPMESVPLPGI 180
DB 121 IYAPFKSLPMPNPDIVFGMDISNMNLADAMARAKVLDIDLOKQLRPMESVPLPGI 180

```
|||||
Db IYAFKSLIPVNPDDIVFGGWDISNMMLADSMTRAKLIDLOKOLRPYMESWVPLPGI 180
121
181 YDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 240
181 YDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 240
241 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 300
241 AGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 300
301 LIIGDDFFSGGOTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPOTFRSKEISKS 360
301 LIIGDDFFSGGOTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPOTFRSKEISKS 360
361 NVVDDMVNSNALIYEPGEHPDHVVVYIKVYPVGDSCRAMDEYTSSEIFMGKNTIYLHNTC 420
361 NVVDDMVNSNALIYEPGEHPDHVVVYIKVYPVGDSCRAMDEYTSSEIFMGKNTIYLHNTC 420
421 EDLSLAAPITLIDVLVLAELSTRIOFKAENEGKFHSFHPVATILSYLTAKAPLVPPGTPV 480
421 EDLSLAAPITLIDVLVLAELSTRIOFKAENEGKFHSFHPVATILSYLTAKAPLVPPGTPV 480
481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
481 ALAKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 7

```
T05017
inositol-3-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thaliana
N:Alternate names: protein T19P19.190
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05017
R:Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdemench, P.; Hohelsel, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15394
A:Accession: T05017
A:Molecule type: DNA
A:Residues: 1-511 <BEV>
A:Cross-references: UNIPROT:P42801; EMBL:AL022605
A:Experimental source: cultivar Columbia; BAC clone T19P19
C:Genetics:
A:Map position: 4
A:Introns: 64/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3
A>Note: T19P19.190
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD
```

```
Query Match 89.3%; Score 2349.5; DB 2; Length 511;
Best Local Similarity 87.9%; Pred. No. 3.6e-150;
Matches 449; Conservative 32; Mismatches 29; Indels 1; Gaps 1;

Qy 1 MFENFKVESPNVXYTETETIQSVNYETTELVHENR-NGTYQWIVPKSVNYQFKTNTHV 59
Db 1 MFISBFKVESPVKXTENHISVYDYETTELVEHKTVNGTYQWIVPKSVNYQFKTNTHV 60
60 PRLGMLVGMGNNSTLTGCIYANREDISWATKTKIQOANYFGSLTQASAIRVSGFGE 119
60 PRLGMLVGMGNNSTLTGCIYANREDISWATKTKIQOANYFGSLTQASAIRVSGFGE 119
61 PRLGMLVGLGNNSTLTGCIYANKEGISWATKTKIQOANYFGSLTQASAIRVSGFGE 120
61 PRLGMLVGLGNNSTLTGCIYANKEGISWATKTKIQOANYFGSLTQASAIRVSGFGE 120
Qy 120 EIVAPFKSLIPVNPDDIVFGGWDISNMMLADSMTRAKLIDLOKOLRPYMESWVPLPG 179
Db 121 EIVAPFKSLIPVNPDDIVFGGWDISNMMLADSMTRAKLIDLOKOLRPYMESWVPLPG 180
180 IYDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 239
180 IYDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 239
181 IFDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 240
181 IFDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 240
240 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 299
240 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 299
241 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 300
241 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 300
```

```
Qy 300 TLIGDDFFSGGOTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPOTFRSKEISKS 359
Db 301 VLIIGDDFFSGGOTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPOTFRSKEISKS 360
Qy 360 SNVDDMVNSNALIYEPGEHPDHVVVYIKVYPVGDSCRAMDEYTSSEIFMGKNTIYLHNT 419
Db 361 SNVDDMVNSNALIYEPGEHPDHVVVYIKVYPVGDSCRAMDEYTSSEIFMGKNTIYLHNT 420
Qy 420 CEDSLAAPITLIDVLVLAELSTRIOFKAENEGKFHSFHPVATILSYLTAKAPLVPPGTPV 479
Db 421 CEDSLAAPITLIDVLVLAELSTRIOFKAENEGKFHSFHPVATILSYLTAKAPLVPPGTPV 480
Qy 480 NALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 NALSKORAMLENIMRACVGLAPENNMILEYK 511
```

RESULT 8

```
T04339
inositol-3-phosphate synthase (EC 5.5.1.4) - barley
C:Species: Hordeum vulgare (barley)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04339
R:Larson, S.R.; Raboy, V.
submitted to the EMBL Data Library, March 1998
A:Description: linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A:Reference number: Z14366
A:Accession: T04339
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-510 <LAR>
A:Cross-references: UNIPROT:O65195; EMBL:AF056325; NID:g3152730; PIDN:AACT17133.1; PID:g31
A:Experimental source: cv. Harrington
C:Genetics:
A:Gene: INO1
A:Map position: 4
A:Function:
A:Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inoc
A:Pathway: inositol biosynthesis
A>Note: first step
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD
```

```
Query Match 88.9%; Score 2338; DB 2; Length 510;
Best Local Similarity 88.0%; Pred. No. 2.1e-149;
Matches 449; Conservative 26; Mismatches 35; Indels 0; Gaps 0;

Qy 1 MFENFKVESPNVXYTETETIQSVNYETTELVHENRNGTYQWIVPKSVNYQFKTNTHV 60
Db 1 MFISBFKVESPVKXTENHISVYDYETTELVEHSHDGASKVVRPKSVNYHFKTNTHV 60
61 PRLGMLVGMGNNSTLTGCIYANREDISWATKTKIQOANYFGSLTQASAIRVSGFGE 120
61 PRLGMLVGMGNNSTLTGCIYANREDISWATKTKIQOANYFGSLTQASAIRVSGFGE 120
Qy 121 IYAFKSLIPVNPDDIVFGGWDISNMMLADSMTRAKLIDLOKOLRPYMESWVPLPGI 180
Db 121 IYAFKSLIPVNPDDIVFGGWDISNMMLADSMTRAKLIDLOKOLRPYMESWVPLPGI 180
181 YDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 240
181 YDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 240
181 YDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 240
181 YDPDFIANDGERANVYIKGTKEQVOOIIKDIAFKREATVDKVVYVMTANTERYSNLV 240
241 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 300
241 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 300
241 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 300
241 VGLNDTMENTLAAVDRNEAEISPSITLYAIACWENVPPINGSPONTFVPGIIDLAIANT 300
301 LIIGDDFFSGGOTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPOTFRSKEISKS 360
301 LIIGDDFFSGGOTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPOTFRSKEISKS 360
301 LIIGDDFFSGGOTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPOTFRSKEISKS 360
301 LIIGDDFFSGGOTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNMLSAPOTFRSKEISKS 360
361 NVVDDMVNSNALIYEPGEHPDHVVVYIKVYPVGDSCRAMDEYTSSEIFMGKNTIYLHNTC 420
361 NVVDDMVNSNALIYEPGEHPDHVVVYIKVYPVGDSCRAMDEYTSSEIFMGKNTIYLHNTC 420
```


A:Residues: 1-525 <W12>
A:Cross-references: UNIPROT:Q16664; EMBL:AL033535; PDB:CAA22132.1; CESP:VF13D12L.1
A:Experimental source: clone VF13D12L
R:Gajdety, S.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19209
A:Accession: I20002
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-525 <W12>
A:Cross-references: EMBL:Z69902; PDB:CAA93771.1; GSPDB:GN00020; CESP:VF13D12L.1
A:Experimental source: clone CA7D12
C:Genetics:
A:Gene: CESP:VF13D12L.1
A:Map position: 2
A:Introns: 106/2; 287/1; 411/2
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 53.0%; Score 1393.5; DB 2; Length 525;
Best Local Similarity 53.3%, Pred. No. 6.4e-86;
Matches 269; Conservative 93; Mismatches 134; Indels 9; Gaps 5;

Dy VESPNVAKTEEISOVYNETTELVEHNRNTQOMIVKPKSVNQFKTTHVPLGLVMVL 67
||| | :
Db VESPNVKLEDVLSRFYRKHFH-RADGLH---VTPKHDSFRTVLKPRKTGLLV 70
68 GWGGNGSTLTGVIANREDISMATKDKIQQANYFGSLTQAASIRVG--SFQGEIYAPF 125
Qy ||| | :
Db GLGGNGSTANGSIFANQYAMTWRTKEHSQANFYGSVTQTATVHLGDASTQNQIVPF 130
Qy 126 KSLPMPNVDPDIVFGGWDISNMTADAMARKVPDIDLOKOLRYMESWVLPPIYPDF 185
Dy :
131 KDIVPILSPNDLIISGMDISNSNLYEAMGRAKVEPELQEKLRPMERIVPLPSIYYPDF 190
Qy 186 IAAOGEERANNVIKG-TKEQVOOIKDIKAFKTKDKVYVMTANTREYSNLVGLN 244
Db ||| | :
191 IASNOGDANNVIPEDNKLEHLEHRADIRKFQKSHHECVYIMTANTEHYTDVRGLN 250
Qy 245 DTMENLLAANDNEAEISPSTLYAIACMVENVPIINGSPONTFVPGHIDLIAENTLIG 304
Db :
251 ATADEIMSIINVNDDEVSPSNIPAVASILBEGAHYNOSPQTLLVPGIIELEARKHVVG 310
Qy 305 DDFFSGGTOKMSVLYDLFVAGIKPTSYVNHGNNDGNLSPQTFRSKEISKSVND 364
Db :
311 DDFFSGGTOKFSAAFDFVSSGMRPESIVSYNHGNNDGNLSEAROFRSKEISKSVND 370
Qy 365 DMVNSMILYEPGHPPDHVVIVKYVPVYGSKRAMDETSRIFMGKNITVILHTGCDSL 424
Db ||| | :
371 DMVNSNQLLFDPADKNPDYCVVIKVVPYADSKRAMDETSCI FMGKQTFVHNHCBSL 430
Qy 425 LAAPILIIDLVAELSTRIOFAENEGKFHFPHAVATILSYLAAPLVPCTPVNALSK 484
Db 431 LASPLIYLALLTELASVSXKVDDE--YKRFHSVLSLISLLRLRPVPROPIISNFMR 488
Qy 485 QRAMLLENIMRAQVGIAPENNMLEY 509
Dy ||| | :
489 QESTLTKLVTLAAGPPSDTDWQIER 513

RESULT 12
S45452 inositol-3-phosphate synthase (EC 5.5.1.4) - yeast (Candida albicans)
C:Species: Candida albicans
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S45452
R:Kilig, L.S.; Zobel, P.A.; Devry, C.G.; Losberger, C.
yeast 10, 789-800, 1994
A>Title: Comparison of INO1 gene sequences and products in Candida albicans and Saccharomyces cerevisiae
A:Reference number: S45452; MUID:95066381; PMID:7975896
A:Accession: S45452
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA

A:Residues: 1-520 <K1>
A:Cross-references: UNIPROT:P42800; EMBL:L22737; NID:g413750; PDB:AAA62849.1; PID:g69575
C:Genetic8:
C:Gene: INO1
C:Function:
A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-
A>Note: requires NAD
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 52.8%; Score 1389.5; DB 2; Length 520;
Best Local Similarity 52.1%; Pred. No. 1.2e-85;
Matches 267; Conservative 98; Mismatches 132; Indels 15; Gaps 5;

Oy ESPVAKYETETIQSYNYNTEETTELVAHNRNGTQWLYKPSKSNVQEFKTNHVPFLGLVWLVG 68
Db KSSSVTDHDLTKFTYENS-VVEKDANG-KFIYTPASDIEFFYVDLPPVYGLLVG 64
Oy WGNNGSTLTGTVANREDISWATKDKIOQANVPFSLTQASARVVG-SFOGEIYAPFK 126
Db IGGNNGITLLGATLTKNKNISFENKEGVKPKYGSVTQASTYKIGVDKETGSDVYVFN 124
Oy SLLEPVNEDDIYFGGMDISNNMLADAMARAKYFDIDLQKOLRPMESMVPFLPGIYDFDI 186
Db SLVPPVNNNDLVQDMDISGLPLDQAMKRAKVLDTLQKQVLYLENKKPLRESIYDFDI 184
Oy 187 AANDEERANNVYK-----GTQEOVQOITIKDIAFKFPAKATVDKVVVLTMTANTERYSN 238
Db 185 ALNOSERANNVFNQVNGEYKTDINKNADYEKIKRDIRDFPAKAKELDKVITILMTANTERYAD 244
Oy 239 LVVGLINDTMENTLAAVDNEAEISPTLTAICWENVPFINGSPONTFVPGILDLAIAR 298
Db 245 VLPVNDTPADNLTIKSIKSHERIASVFAVASILTKVYIINSPOITFVPGVIELKEXY 304
Oy 299 NTLIGDDPFKSGQTQAKSVLVDPLVAGAGIKPISIVSYNHLGNNDGNMLSAQTFRSKEIS 358
Db 305 DSFLIGDDPFKSGQTQAKSVLVDPLVAGAGIKPISIVSYNHLGNNDGNMLSAQTFRSKEIS 364
Oy 359 KSNVVDNVSNNALLY--EPGEHPHVVIKTVPYVGDGSRAMDEYTSSEIFMGKNTIVL 416
Db 365 KQSVVDITIESENELLYNKGSGDKVDHCIVIKYLPVAGDCKVAMDEYSELMGLGHNKISI 424
Oy 417 HNTGDSLLAPILIDLVLAELSTRIOFKANEGGFHFPATLISYLTKAPLVPGT 476
Db 425 HNVGDSLLAPILIDLVVALEFATRVQVKGKSDYDELVPVASLISLTWKAPLARPGE 484
Oy 477 PVVNAALSKORAMLEINRACVGLAPENNMLE 508
Db 485 KPINGANKROQLVNLISVGLPIDNELRFE 516

RESULT 13
A30902
Inositol-3-phosphate synthase (EC 5.5.1.4) [validated] - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein U0610; protein YOL153c
C:Species: Saccharomyces cerevisiae
C:Date: 18-Apr-1989 #sequence revision 08-Sep-1995 #text change 16-Aug-2004
C:Accession: S55160; B32209; S55935; S71644; A30877; A30902
R:Karsoulou, C.; Tzermla, M.; Alexandraki, D.
A:Submitted to the EMBL Data Library, May 1995
A:Description: The complete sequence of a 40.7 kb segment located on the left arm of yeast
yeast hypochetrical proteins.
A:Reference number: S55159
A:Accession: S55160
A:Molecule type: DNA
A:Residues: 1-555 <KAT>
A:Cross-references: EMBL:X87371; NID:g854542; PID:g854544
R:Dean-Johnson, M.; Henry, S.A.
J: Biol. Chem. 264, 1274-1283, 1989
A:Title: Biosynthesis of inositol in yeast. Primary structure of myo-inositol-1-phosphate
A:Reference number: A32209; MUID:89093118; PMID:2642502
A:Accession: B32209
A:Molecule type: DNA

A:Residues: 23-35, 'RL', 37-81, 'FE', 83-87, 'TRNYAHVWV', 88, 'QQW', 92-103, 'WPKYRISTMTS', 116-
VISFORLSFSFSAVY', <DEA>
A:Cross-references: EMBL:J04453
R.Katsoulou, C.; Tsermla, M.; Alexandraki, D.
Submitted to the Protein Sequence Database, September 1995
A:Reference number: S56912
A:Accession: S56935
A:Molecule type: DNA
A:Residues: 1-555 <KAW>
A:Cross-references: EMBL:Z49428; NID:g1015570; PID:g1015571; MIPS:YJL153c
R.Katsoulou, C.; Tsermla, M.; Tavernarakis, N.; Alexandraki, D.
Yeast 12, 787-797, 1996
A:Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X
of chromosome XI.
A:Reference number: S71643; MUID:96408771; PMID:8813765
A:Accession: S71644
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-555 <KAP>
A:Cross-references: EMBL:X87371; NID:g854542; PID:CAA60802.1; PID:g854544
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
A:Gene: SGD:INO1
A:Cross-references: SGD:S0003689; MIPS:YJL153c
A:Map position: 10L
C:Complex: homotetramer
C:Function: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo
A:Pathway: inositol biosynthesis
A>Note: requires NAD
C:Superfamily: Myo-inositol-1-phosphate synthase
C:Keywords: cytosol; homotetramer; intramolecular lyase; isomerase; NAD

Query Match 52.8%; Score 1389.5; DB 2; Length 555;
Best Local Similarity 53.1%; Pred. No. 1.3e-85;
Matches 278; Conservative 87; Mismatches 138; Indels 21; Gaps 8;

QY 3 IENFVSEPNKYTETETIQSVYNYETETELVHNRGTYQWYIKPSVNYQFNTNHP-K 61
DB 31 ITSVAVVDKCTYKQNELTKYSENA-VVTKTASGRD-VTPVQDVDFLDLKKBEK 87
QY 62 LGVNLVWGCGNNGSTLTGIVANREDISWATKDIQANFYGLTQASAIRVG-SFOGEE 120
DB 88 LGIMLIGCGNNGSTLVASVLANKNVETQTEGKQRYFESMTQCSTLKIGIABGND 147
QY 121 IYAPFKSLPWNPDIVFGGWDISNMNLADAMAKYFDIDLOQLRPYMSWVPLRGI 180
DB 148 IYAPFNSTLPMVSPVDSGWDINNADLYEAMQSRQVLEVDLOQLRKAQSLVAKPLPSI 207
QY 181 YDPDPIANQEBRANV-----KGT-----KQEQVQIITKIKAKREKATKDKVVLMTA 231
DB 208 YYPDPIANQEBRANNCINLDEKGVNTRGKWTNLIQRIIRDIONFKENALDKVILMTA 267
QY 232 NTERSNLWGLNDMENLILAAVDNBEAISPSTLYAICWENYPTNGSPONTFVPGI 291
DB 268 NTERVAVSPGVNDIMENLLOS IKNDHEIAPSTIFLAASILEGVPTNGSPONTFVPGI 327
QY 292 IDLAIAARNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNGMNLAPOT 351
DB 328 VOLAHEGTFIAGDDLKSGQTKMSVLVLAQFLVDAGIKPTSTIVSYNHLGNNGMNLAPOT 387
QY 352 FRSKKISKSNVVDVMSNNAIYER-GEHPHVIVYIKVYVPGDSKRAMDEYTSIEIMG 409
DB 388 FRSKKISKSNVVDVMSNNAIYER-GEHPHVIVYIKVYVPGDSKRAMDEYTSIEIMG 447
QY 410 GKNTIVLHNTCEDSLAIPILIDVLVLAELSTRIOFK-----AENEGKPHSPVATILS 464
DB 448 GHNRSIHNVCEDSLALPILIDVLVLAELSTRIOFK-----AENEGKPHSPVATILS 507
QY 465 YLTAAVLPVPGTVPVVALSKOPAMLEIMRAQVGLAPENNMLE 508
DB 508 YMLKAPLTPRGFHPVNGINKQRTALENFLRLILIGPQNEIRAFE 551

RESULT 14
146317
hypothetical protein DKFZp434A0612.1 - human
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
R:Accession: T46317
Submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23035
A:Accession: T46317
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-430 <AAA>
A:Cross-references: UNIPROT:Q9NSU0; EMBL:AL137749
A:Experimental source: adult testis; clone DKFZp434A0612
C:Genetics:
A>Note: DKFZp434A0612.1
C:Superfamily: myo-inositol-1-phosphate synthase

Query Match 49.9%; Score 1312.5; DB 2; Length 430;
Best Local Similarity 64.2%; Pred. No. 1.3e-80;
Matches 244; Conservative 66; Mismatches 69; Indels 1; Gaps 1;

QY 131 MWNPDIVFGGWDISNMNLADAMAKYFDIDLOQLRPYMSWVPLRGIYDPDPIANQ 190
DB 1 MVAEPDILVFDGWDISSILALAEAMRAKVLWDGLOLMPHMEALPRPSVYIPEFIANQ 60
QY 191 EERANNVKTGKQEQVOQIINDIKAFKATKDKVAVLMTANTERYSNLVGLANTMENT 250
DB 61 SARBDNLPGRSQQLBEIRDIRNFRSSAGLDKIVLMTNTEFCVIRIPLANTANT 120
QY 251 LAVDNREAEISPSTLYAICWENYPTNGSPONTFVPGIIDLAIARNTLIGDDPKSG 310
DB 121 IRTTELGL-LEVSPTLFAVASILBEGCAFINGSPQITLVPGALIELAMQHRVFGGDDPKSG 179
QY 311 QTKMSVYVDFLVAGIKPTSTIVSYNHLGNNGMNLAPOTFRSKISKSNVVDVMSN 370
DB 180 QTKMSVYVDFLVAGIKPTSTIVSYNHLGNNGMNLAPOTFRSKISKSNVVDVMSN 239
QY 371 AILVEGEHPHVIVYIKVYVPGDSKRAMDEYTSIEIPMGKNTIYVHNTCEDSLAIP 430
DB 240 PVLITRGSEPHCVIKVYVPGDSKRAMDEYTSIEIPMGKNTIYVHNTCEDSLAIP 299
QY 431 IDVLVLAELSTRIOFKAENEGKPHSPVATILSYTAPLVPVPGTVPVVALSKORAMLE 490
DB 300 IDLALITELCQVAFCTMDPEPQTFHPVLSLSFLFAPLVPVPGTVPVVALSKORAMLE 359
QY 491 NIMRACVGLAPENNMLEYK 510
DB 360 NIMRACVGLAPENNMLEYK 379

RESULT 15
T36191
problee myo-inositol phosphate synthase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
R:Accession: T36191
Submitted to the EMBL Data Library, March 1999
A:Reference number: Z21600
A:Accession: T36191
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-388 <SNU>
A:Cross-references: UNIPROT:Q9Z4Y8; EMBL:AL035707; PIDN:CA838887.1; GSPDB:GN00070; SCOECD
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOECD:SCE29.12C
C:Superfamily: myo-inositol-1-phosphate synthase

Query Match 11.1%; Score 291; DB 2; Length 388;

Best Local Similarity 27.7%; Pred. No. 4.8e-12;
Matches 110; Conservative 61; Mismatches 158; Indels 68; Gaps 11;

```

QY      61  KLGWLVGMGNNSTLGGVIANREDISWATKDIQGANFGLTQASAIRVGSFQGE 120
      Db      2  RTGIMLVGARGSVATTYVGAALRAGLVPAI-----GCVSALRA----FDG-- 44
QY      121  IYAPFKSLPMVNPDDIVFGMDISNMNI---ADAMARAKYFDIDLQKLRPYMESWVP- 176
      Db      45  -----VPLPGFDELVFGGHDVGTGLVKRAEQLAEGV-----VPR 80
QY      177  -LPGIYDPDFIAANOEEERANNVIKTKQ-----EQVOIIRKIKAFKEATKVDKV 226
      Db      81  GLPGVLTAEIDAADAIEIRP--APGTEEGGDGTPDQASAGAIVADLTGFRERLGLDRVV 138
QY      227  VLMWTANFERVSNLVVGLNDTMENTLLAAVDRNEABISPSLYALACVMENVPFINGSPQ-N 285
      Db      139  VVNVSSITPPA-VPHPAHASAAALREALARGERPLPVSSLYAYALRAGCAFDFTPSTG 197
QY      286  TFVPGQLDLATARTLLIGDDFKSGQTKKSVLVDFLVGAGIKPTSIYSYNHLGNNDGMN 345
      Db      198  ARLPALDELAREQGLPYAGSDGKTGETLVKSVLAFMFARBALRVRSWSGTNLLGGGDGAT 257
QY      346  LSAPOTFRSKEISKSNVVDVMVNSNAIILYEPGEHPDHVVVIKYVPYVGDSTRANDDEYTS 405
      Db      258  LADPERVVSKNASKGLVLE-----AELGHAVEGGVHIHVVDLGEWKTAMDHYTFE 308
QY      406  IFMGKNTIVLHNTCEDSLAAPTILDLVLAELSTR 442
      Db      309  GFLGARMTLQFTWQCGDSSLAAPLVLDLARFMALNHR 345

```

Search completed: June 7, 2005, 16:47:42
Job time : 26.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 15:09:19 ; Search time 92.5 Seconds
(without alignment)

2132.409 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631
Sequence: 1 MFENFKFESPNVKTETETI.....NIMRACVGLAPBNMILEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 366760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2631	100.0	510	ADQ14501	Adq14501 Mutant bo
2	2631	100.0	510	ADS82002	Adsg82002 Soybean m
3	2624	99.7	510	ADQ14503	Adq14503 Mutant bo
4	2624	99.7	510	ADQ14505	Adq14505 Wild type
5	2624	99.7	510	ADS82004	Adsg82004 Soybean m
6	2624	99.7	510	ADS82006	Adsg82006 Soybean m
7	2593	98.6	510	AAW79740	Aaw79740 Soybean w
8	2593	98.6	510	ADQ14499	Adq14499 Mutant bo
9	2593	98.6	510	ADQ14491	Adq14491 Wild type
10	2593	98.6	510	ADS82000	Adsg82000 Soybean m
11	2593	98.6	510	ADS81994	Adsg81994 Soybean m
12	2588	98.4	510	AAW79741	Aaw79741 Soybean m
13	2588	98.4	510	ADQ14495	Adq14495 Mutant bo
14	2588	98.4	510	ADS81998	Adsg81998 Soybean m
15	2470	93.9	536	AAW24477	Aaw24477 Nicotiana
16	2431	92.4	505	ADG73738	Adg73738 Myo-inosi
17	2384	90.6	510	AAAG50575	AAg50575 Arabidops
18	2384	90.6	510	ADW73525	Adw73525 Thale cre
19	2384	90.6	581	AAAG50574	AAg50574 Arabidops
20	2384	90.6	645	AAAG50573	AAg50573 Arabidops
21	2376	90.3	510	AAAB48935	AAb48935 Brassica
22	2354	89.3	510	AAW96259	Aaw96259 Phytate p
23	2349.5	89.3	511	AAAG09861	AAg09861 Arabidops
24	2349.5	89.3	511	AAAG32501	AAg32501 Arabidops
25	2349.5	89.3	534	AAAG09860	AAg09860 Arabidops

26	2346	89.2	510	2	AAW97882	Aaw97882 Maize myo
27	2345	89.1	510	8	ADP43920	Adp43920 Rice L-my
28	2337	88.8	510	4	AAAB47286	AAb47286 MIP synth
29	2084	79.2	446	3	AAAG32502	AAg32502 Arabidops
30	2084	79.2	446	3	AAAG09862	AAg09862 Arabidops
31	1883	71.6	512	8	ADP43919	Adp43919 Porterei
32	1779	67.6	380	3	AAAG32503	AAg32503 Arabidops
33	1616	61.4	558	4	AAAG65577	AAg65577 Human hml
34	1616	61.4	558	4	AAAM40076	AAm40076 Human pol
35	1616	61.4	558	4	AAAB92849	AAb92849 Human pro
36	1616	61.4	558	4	AAAB93732	AAb93732 Human pro
37	1616	61.4	538	5	AAO18891	AAo18891 Human ova
38	1612.5	61.3	534	6	ABJ25857	Abj25857 Aspergill
39	1608.5	61.1	534	6	ABJ26457	Abj26457 Aspergill
40	1602	60.9	557	4	AAAG65576	AAg65576 Murine m
41	1602	60.9	565	4	ABAB65497	ABb65497 Drosophi
42	1565.5	59.5	659	8	ADH99714	Adh99714 Novel hum
43	1538	58.5	505	8	ADG73737	Adg73737 Aspergill
44	1538	58.5	532	8	ADG73736	Adg73736 Aspergill
45	1429	54.3	504	7	ADB64735	ADB64735 Human pro

ALIGNMENTS

RESULT 1	ADQ14501	ADQ14501 standard; protein; 510 AA.
XX	ADQ14501;	
AC	ADQ14501;	
XX	23-SEP-2004 (first entry)	
DT		
XX		
DE	Mutant soybean myo-inositol 1-phosphate synthase polypeptide #3.	
XX		
KW	Soybean: myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;	
KW	raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;	
KW	mutant; mulein.	
OS	Glycine max.	
OS	Synthetic.	
PN	US2004128713-A1.	
XX		
PD	01-JUL-2004.	
XX		
PF	21-NOV-2003; 2003US-00718952.	
XX		
PR	08-APR-1997; 97US-00835751.	
PR	07-APR-1998; 98WO-US006822.	
PR	26-APR-1999; 99US-00299315.	
XX	11-MAR-2002; 2002US-00025003.	
XX		
PA	(HITZ/) HITZ W D.	
PA	(SEBA/) SEBASTIAN S A.	
PA	(GRAC/) GRACE D J.	
XX	(STRE/) STREIT L G.	
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
XX	WPI; 2004-53135/51.	
DR	N-Psdb; ADQ14500.	
XX		
PT	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,	
PT	useful for producing plants with decreased raffinose, stachyose, and	
PT	phytic acid and increased sucrose, leading to valuable and useful soybean	
PT	products.	
XX		
PS	Claim 9; SEQ ID NO 12; 48pp; English.	
XX		
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-	
CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate	
CC	synthase having a decreased capacity for the synthesis of myo-inositol 1-	

CC phosphatase. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

XX

Sequence	510 AA;
Query Match	100.0%; Score 2631; DB 8; Length 510;
Best Local Similarity	100.0%; Pred. No. 8.7e-215;
Matches 510; Conservative 0;	Mismatches 0; Indels 0; Gaps 0

[illegible]

OS	Glycine max; line 29010CPR1.
XS	Synthetic.
XX	
XX	Key
XX	Location/Qualifiers
XX	Misc-difference 87
XX	/note= "Wild-type Gly substituted by Asp"
XX	
XX	US2003074685-A1.
XX	
XX	17-APR-2003.
XX	
XX	11-MAR-2002; 2002US-00025003.
XX	
XX	08-APR-1997; 97US-00835751.
XX	07-APR-1998; 98WO-US006822.
XX	
XX	(HITZ/) HITZ W D.
XX	(SEBA/) SEBASTIAN S A.
XX	
XX	Hitz WD, Sebaetian SA;
XX	
XX	WPI; 2004-639957/62.
XX	
XX	N-PSDB; ADS82001.
XX	
XX	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX	
XX	Claim 9; SEQ ID NO 12; 34pp; English.

				The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not RR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant mentioned above), seeds crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products. Since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a mutant myo-inositol 1-phosphate synthase.
XX	Sequence 510 AA;			
SO				
	Query Match	100.0%;	Score 2631;	DB #6; Length 510;
	Best Local Similarity	100.0%;	Pred. No. 8.7e-216;	
	Matches 510;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0
OY	1 MFIEIFKESNVNKKTERETIISVNVYFETTELVEHNRNGYQWIVPKSVNYOFKNTHTVP	60		
Ds	1 MFIEIFKESPVNKKTERETIISVNVYFETTELVEHNRNGTQWIVPKSNVYFKNTHTVP	60		
OY	KLGVLGVGMGNNGGTLTGCVIANREDISWATKDKIQQANFYGSLLTOASAIIVGSFGEE	120		

Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKXIQOANYFGSLTQASAIRVSGFOEE 120
 QY 121 IYAPFKSLIPWNPDDIVFGGWDISNMNLADAMAAKVPDIDLOKQALRPYHESWVPLGCI 180
 Db 121 IYAPFKSLIPWNPDDIVFGGWDISNMNLADAMAAKVPDIDLOKQALRPYHESWVPLGCI 180
 QY 181 YDPDFIANDOEERANNVIKGTQOEVOQI IKDIAFKKATKVDKVVVMTANTERYSNLV 240
 Db 181 YDPDFIANDOEERANNVIKGTQOEVOQI IKDIAFKKATKVDKVVVMTANTERYSNLV 240
 QY 241 VGLNDTMENLLAAVDNRNEAISPSTLYAIACMENVPFINSPOPTFVPGGLIDLAIAANT 300
 Db 241 VGLNDTMENLLAAVDNRNEAISPSTLYAIACMENVPFINSPOPTFVPGGLIDLAIAANT 300
 QY 301 LIGGDDFRSGGQTKKMSVVDVFLVGAGIKPTSVSYNHLGNDGNLSPQTFRSKEISKS 360
 Db 301 LIGGDDFRSGGQTKKMSVVDVFLVGAGIKPTSVSYNHLGNDGNLSPQTFRSKEISKS 360
 QY 361 NVVDDMVNSNAILYEPGEHPDHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
 Db 361 NVVDDMVNSNAILYEPGEHPDHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
 QY 421 EDSLILAAPIIIDVLVLAELSTRIOFKAENEGKFSFHPVATILSYLTRAPLVPFGTPIVYN 480
 Db 421 EDSLILAAPIIIDVLVLAELSTRIOFKAENEGKFSFHPVATILSYLTRAPLVPFGTPIVYN 480
 QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
 Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 3
 ID ADO14503 standard; protein; 510 AA.
 AC ADO14503;

DT 23-SEP-2004 (first entry)
 XX Mutant soybean myo-inositol 1-phosphate synthase polypeptide #4.
 DE Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
 KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
 KM mutant; mtein.
 XX Glycine max.
 OS Synthetic.
 XX US2004128713-A1.
 PN 01-JUL-2004.
 XX 21-NOV-2003; 2003US-00718952.
 PF 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-00293315.
 PR 11-MAR-2002; 2002US-00025003.
 XX (HITZ/) HITZ M D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 XX WPI: 2004-533135/51.
 DR N-PSDB; ADO14502.
 XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.

XX Example 8; SEQ ID NO 14; 48pp; English.

PS The invention relates to a nucleic acid fragment encoding a soybean myo-
 XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant, a
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents a mutant soybean myo-inositol
 CC 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 99.7%; Score 2624; DB 8; Length 510;
 Best Local Similarity 99.8%; Pred. No. 3.4e-215;
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFIEHFVESPNVXYTETETIQSVYNYETTELVHNRNCTYQMIYKPSVNYQKTNTHVP 60
 Db 1 MFIEHFVESPNVXYTETETIQSVYNYETTELVHNRNCTYQMIYKPSVNYQKTNTHVP 60
 QY 61 KLGWLVGMGNGNSTLTGVIANREDISWATKXIQOANYFGSLTQASAIRVSGFOEE 120
 Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKXIQOANYFGSLTQASAIRVSGFOEE 120
 QY 121 IYAPFKSLIPWNPDDIVFGGWDISNMNLADAMAAKVPDIDLOKQALRPYHESWVPLGCI 180
 Db 121 IYAPFKSLIPWNPDDIVFGGWDISNMNLADAMAAKVPDIDLOKQALRPYHESWVPLGCI 180
 QY 181 YDPDFIANDOEERANNVIKGTQOEVOQI IKDIAFKKATKVDKVVVMTANTERYSNLV 240
 Db 181 YDPDFIANDOEERANNVIKGTQOEVOQI IKDIAFKKATKVDKVVVMTANTERYSNLV 240
 QY 241 VGLNDTMENLLAAVDNRNEAISPSTLYAIACMENVPFINSPOPTFVPGGLIDLAIAANT 300
 Db 241 VGLNDTMENLLAAVDNRNEAISPSTLYAIACMENVPFINSPOPTFVPGGLIDLAIAANT 300
 QY 301 LIGGDDFRSGGQTKKMSVVDVFLVGAGIKPTSVSYNHLGNDGNLSPQTFRSKEISKS 360
 Db 301 LIGGDDFRSGGQTKKMSVVDVFLVGAGIKPTSVSYNHLGNDGNLSPQTFRSKEISKS 360
 QY 361 NVVDDMVNSNAILYEPGEHPDHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
 Db 361 NVVDDMVNSNAILYEPGEHPDHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
 QY 421 EDSLILAAPIIIDVLVLAELSTRIOFKAENEGKFSFHPVATILSYLTRAPLVPFGTPIVYN 480
 Db 421 EDSLILAAPIIIDVLVLAELSTRIOFKAENEGKFSFHPVATILSYLTRAPLVPFGTPIVYN 480
 QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
 Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 4
 ID ADO14505 standard; protein; 510 AA.
 AC ADO14505;

DT 23-SEP-2004 (first entry)

DE Wild type soybean myo-inositol 1-phosphate synthase polypeptide #2.
XX
KM Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.
OS
XX Glycine max.
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hiltz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-53135/51.
DR N-PSDB; ADQ14504.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Claim 3; SEQ ID NO 16; 48pp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a wild type soybean myo-
CC inositol 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 510 AA;
SQ
Query Match 99.7%; Score 2624; DB 8; Length 510;
Best Local Similarity 99.8%; Pred. No. 3.4e-215;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 181 YDPPFIANOEBERANNVIGKTKOQVOQIIKDIAKFAKATKVKVVLMTANTERYSNLV 240
QY 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVPFINGSPONTFVPGIDIAIRNT 300
DB 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWENVPFINGSPONTFVPGIDIAIRNT 300
QY 301 LIGGDDFKSGQTKKSVLYVDLFGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKKSVLYVDLFGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSKEISKS 360
QY 361 NVYDDMNSNAIILYEPBEHPDHVVVITKVPYVGDGSKRAMDEYSEIFMGKNITVLHNTC 420
DB 361 NVYDDMNSNAIILYEPBEHPDHVVVITKVPYVGDGSKRAMDEYSEIFMGKNITVLHNTC 420
QY 421 EDSLAAPIITLDVLAAELSTRIQFKAENEGKPHSPHPVATIIISYLTKAPLVPGTPVNV 480
DB 421 EDSLAAPIITLDVLAAELSTRIQFKAENEGKPHSPHPVATIIISYLTKAPLVPGTPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
RESULT 5
ADSS82004
ID ADSS82004 standard; protein; 510 AA.
XX
XX ADSS82004;
AC
XX 18-NOV-2004 (first entry)
DT
XX
XX Soybean myo-inositol 1-phosphate synthase wild-type 4.
DE
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29018JP03.
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
PD
PF 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PI Hiltz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
DR N-PSDB; ADS82003.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Example 8; SEQ ID NO 14; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol 1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus

CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homologous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 99.7%; Score 2624; DB 8; Length 510;

Best Local Similarity 99.8%; Pred. No. 3,4e-215;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFIENFKVSPNVKKTETETISQVYNYETTELVEHNRNGTYOMIYKPSKVNQFKNTHTVP 60
DB 1 MFIENFKVSPNVKKTETETISQVYNYETTELVEHNRNGTYOMIYKPSKVNQFKNTHTVP 60
QY 61 KLGVWLVMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGVWLVMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLPVMNPPDIVFGGWDISNMMLADAMARAKVFDIDLQKQLRPYMSWVPLPGI 180
DB 121 IYAPFKSLPVMNPPDIVFGGWDISNMMLADAMARAKVFDIDLQKQLRPYMSWVPLPGI 180
QY 121 IYAPFKSLPVMNPPDIVFGGWDISNMMLADAMARAKVFDIDLQKQLRPYMSWVPLPGI 180
DB 121 IYAPFKSLPVMNPPDIVFGGWDISNMMLADAMARAKVFDIDLQKQLRPYMSWVPLPGI 180
QY 181 YDPPDIANOEERANNVIKGTQEOVOQI IKDIAKFEKATKVDKVVMTANTERYSNLV 240
DB 181 YDPPDIANOEERANNVIKGTQEOVOQI IKDIAKFEKATKVDKVVMTANTERYSNLV 240
QY 241 VGLNDTMENTLAAVRNEAEISPTLYAIACVMEVPIFNGSPONTFVPGILDLIAANT 300
DB 241 VGLNDTMENTLAAVRNEAEISPTLYAIACVMEVPIFNGSPONTFVPGILDLIAANT 300
QY 301 LIGGDDFKSGQTKMSVLVDPLVGAIGIKPTISVYNHLGNDGNMLASAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLVDPLVGAIGIKPTISVYNHLGNDGNMLASAPQTFRSKEISKS 360
QY 361 NVVDPMVNSNALIYEPGHPDHVVYIKYVPYVGSKRAMEYTSSEIFMGKNTIVLANTC 420
DB 361 NVVDPMVNSNALIYEPGHPDHVVYIKYVPYVGSKRAMEYTSSEIFMGKNTIVLANTC 420
QY 421 EDSLLAAPIIIDLVLAELSTRIOFKANEGKESFHPATILSVLTAPLVPPTGPVYN 480
DB 421 EDSLLAAPIIIDLVLAELSTRIOFKANEGKESFHPATILSVLTAPLVPPTGPVYN 480
QY 481 ALSKORAMLENIIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIIMRACVGLAPENNMILEYK 510

RESULT 6
ID ADS82006
ADSS82006 standard; protein; 510 AA.

XX AC ADS82006;

XX DT 18-NOV-2004 (first entry)

XX DE Soybean myo-inositol 1-phosphate synthase wild-type 2.

KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

XX Glycine max; cultivar Wye.

XX US2003074685-A1.

XX 17-APR-2003.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX Hitz WD, Sebastian SA;

XX WPI. 2004-639957/62.

XX N-PSDB; ADS82005.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.

XX Claim 3; SEQ ID NO 16; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreased capacity for the synthesis for myo-
CC inositol 1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homologous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 99.7%; Score 2624; DB 8; Length 510;

Best Local Similarity 99.8%; Pred. No. 3,4e-215;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFIENFKVSPNVKKTETETISQVYNYETTELVEHNRNGTYOMIYKPSKVNQFKNTHTVP 60
DB 1 MFIENFKVSPNVKKTETETISQVYNYETTELVEHNRNGTYOMIYKPSKVNQFKNTHTVP 60
QY 61 KLGVWLVMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGVWLVMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

```
QY 121 IYAPFKSLIPVNPDDIVFGGMDISNMNLADAMARAKVFDDLOKOLRPYMESWVPLPGI 180
DB 121 IYAPFKSLIPVNPDDIVFGGMDISNMNLADAMARAKVFDDLOKOLRPYMESWVPLPGI 180
QY 181 YDPDFIANKOBERANNVIKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERSYSLV 240
DB 181 YDPDFIANKOBERANNVIKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERSYSLV 240
QY 241 VGLNDTMENLLAAVDNRBAEISPSSTLYAIACVMENVPFINGS PONTFVPGIIDLAIANT 300
DB 241 VGLNDTMENLLAAVDNRBAEISPSSTLYAIACVMENVPFINGS PONTFVPGIIDLAIANT 300
QY 301 LIGGDDFKSGQTKKMSVLVDFLVGAGIKPTSIIVSYNHLGNNDGNMLAPQFRSKEISKS 360
DB 301 LIGGDDFKSGQTKKMSVLVDFLVGAGIKPTSIIVSYNHLGNNDGNMLAPQFRSKEISKS 360
QY 361 NVDDDMNSNALILVEPGEHPHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTTVLHNTC 420
DB 361 NVDDDMNSNALILVEPGEHPHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTTVLHNTC 420
QY 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLVPPTGPVNV 480
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLVPPTGPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 7

AAW79740 standard; protein; 510 AA.

AAW79740;

17-OCT-2003 (revised)
02-FEB-1999 (first entry)

Soybean wild-type myo-inositol 1-phosphate synthase.

Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
phytic acid.

Glycine max; line LR13.

MO9845448-A1.

15-OCT-1998.

07-APR-1998; 98WO-US006822.

08-APR-1997; 97US-00835751.

(DUPO) DU PONT DE NEMOURS & CO E I.

Hitz WD, Sebastian SA;

WPI; 1998-568353/48.

N-PSDB; AAV62440.

Soybean plants containing altered myo-inositol-1-phosphate gene - useful
for generating plants with altered levels of e.g. raffinose, stachyose,
phytic acid, etc.

Example 5; Page 45-47; 63pp; English.

This is the amino acid sequence of soybean myo-inositol 1-phosphate
synthase (MI 1-PS) deduced from the coding region of an isolated cDNA
clone (see AAV62440). MI 1-PS is involved in glucose metabolism to phytic
acid, raffinose and stachyose. A mutant MI 1-PS (see AAW79741) has been
identified in soybean line LR13, a mutagenised line of low raffinose
and saccharide phenotype. Sequencing revealed a single base change mutation
in the LR13 gene sequence that resulted in a K396N substitution in the

CC mutant protein. The mutation results in a seed phenotype of very low
CC raffinose saccharide sugars, very high sucrose and low phytic acid. The
CC mutated nucleic acid is used to alter the raffinose saccharide, sucrose,
CC phytic acid and inorganic phosphate content of soybean seeds, leading to
CC useful soybean products, e.g. a seed phytic acid content of less than 17
CC ug/g, a seed content of raffinose and stachyose combined of less than 17
CC 14.5 ug/g, and a seed sucrose content greater than 200 ug/g. (updated on
CC 17-Oct-2003 to standardise OS field)

Sequence 510 AA;

Query Match 98.6%; Score 2593; DB 2; Length 510;
Best Local Similarity 98.4%; Pred. No. 1,5e-212;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 MFENPKVESPNVYKYTEIETISQVNVYETTELVEHNRNGTQOMIVKPKSVNYQKNTHTVP 60
DB 1 MFENPKVECPNVYKYTEIETISQVNVYETTELVEHNRNGTQOMIVKPKSVKTEFNTIHVP 60
QY 61 KLGYMLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVGSFGQEE 120
DB 61 KLGYMLVGMGNGNSTLTGVIANREGISWATKDKIOQANYFGSLTQASAIRVGSFGQEE 120
QY 121 IYAPFKSLIPVNPDDIVFGGMDISNMNLADAMARAKVFDDLOKOLRPYMESWVPLPGI 180
DB 121 IYAPFKSLIPVNPDDIVFGGMDISNMNLADAMARAKVFDDLOKOLRPYMESWVPLPGI 180
QY 181 YDPDFIANKOBERANNVIKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERSYSLV 240
DB 181 YDPDFIANKOBERANNVIKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERSYSLV 240
QY 241 VGLNDTMENLLAAVDNRBAEISPSSTLYAIACVMENVPFINGS PONTFVPGIIDLAIANT 300
DB 241 VGLNDTMENLLAAVDNRBAEISPSSTLYAIACVMENVPFINGS PONTFVPGIIDLAIANT 300
QY 301 LIGGDDFKSGQTKKMSVLVDFLVGAGIKPTSIIVSYNHLGNNDGNMLAPQFRSKEISKS 360
DB 301 LIGGDDFKSGQTKKMSVLVDFLVGAGIKPTSIIVSYNHLGNNDGNMLAPQFRSKEISKS 360
QY 361 NVDDDMNSNALILVEPGEHPHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTTVLHNTC 420
DB 361 NVDDDMNSNALILVEPGEHPHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTTVLHNTC 420
QY 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLVPPTGPVNV 480
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLVPPTGPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 8

ADQ14499 standard; protein; 510 AA.

ADQ14499;

23-SEP-2004 (first entry)

Mutant soybean myo-inositol 1-phosphate synthase polypeptide #2.

Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
mutant; mtein.

Glycine max.

Synthetic.

US2004128713-A1.

01-JUL-2004.

```

PF 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-533135/51.
DR N-PSDB; ADQ14498.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Example 8; SEQ ID NO 10; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents a mutant soybean myo-inositol
XX 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 510 AA;
XX
XX Query Match 98.6%; Score 2593; DB 8; Length 510;
XX Best Local Similarity 98.4%; Pred. No. 1.5e-212;
XX Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 361 NVVDDMVNSNAIIVERGEHPHVVKYVPYVGDGSKRAMDEXTSEIFMGSKNTIVLHNTC 420
DB 361 NVVDDMVNSNAIIVERGEHPHVVKYVPYVGDGSKRAMDEXTSEIFMGSKNTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLLAEISTRIQPKAENEGKFSFHPVATILSYLTRAFLVPPTPVVN 480
DB 421 EDSLAAPIIIDLVLLAEISTRIEKFANEKGFHSFHPVATILSYLTRAFLVPPTPVVN 480
QY 481 ALSKORAMLEINIRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLEINIRACVGLAPENNMILEYK 510
XX
XX RESULT 9
XX ADQ14491
XX ID ADQ14491 standard; protein, 510 AA.
XX
XX ADQ14491;
AC
XX 23-SEP-2004 (first entry)
XX
XX DE Wild type soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX
XX KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
XX raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.
XX
XX OS Glycine max.
XX
XX PN US2004128713-A1.
XX
XX PD 01-JUL-2004.
XX
XX PF 21-NOV-2003; 2003US-00718952.
XX
XX PR 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX PR 26-APR-1999; 99US-00299315.
XX PR 11-MAR-2002; 2002US-00025003.
XX
XX PA (HITZ/) HITZ W D.
XX PA (SEBA/) SEBASTIAN S A.
XX PA (GRAC/) GRACE D J.
XX PA (STRE/) STREIT L G.
XX
XX PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-533135/51.
XX DR N-PSDB; ADQ14490.
XX
XX PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX PS Claim 3; SEQ ID NO 2; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents a wild type soybean myo-

```

CC inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 98.6%; Score 2593; DB 8; Length 510;

Best Local Similarity 98.4%; Pred. No. 1.5e-212; Mismatches 4; Indels 0; Gaps 0;

Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 MFENFKVESPNVKKYETETETISVYNYETTELVEHNRNGTQWIVKPSVNYQFKTNTHVP 60
DB 1 MFENFKVECPNVKKTETETISVYNYETTELVEHNRNGTQWIVKPSVKVEFKTNTHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIYANREDISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIYANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
QY 121 IYAPFKSLPWNPDIVFGGMDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
DB 121 IYAPFKSLPWNPDIVFGGMDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
QY 181 YDPDFIANOEERANNVIKGTKEQVQOIIKDIAFKKATKVDKVVLTANTERYSLV 240
DB 181 YDPDFIANOEERANNVIKGTKEQVQOIIKDIAFKKATKVDKVVLTANTERYSLV 240
QY 241 VGLNDTMENLLAAVDRNEAETSPSTLYAIACMVNVPFINGSPONTFVPGILDAIARNT 300
DB 241 VGLNDTMENLLAAVDRNEAETSPSTLYAIACMVNVPFINGSPONTFVPGILDAIARNT 300
QY 301 LIGGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNNGNLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNNGNLSAPQTFRSKEISKS 360
QY 361 NVVDMDMNSNAIIVYEGEHPHVVIVKVPYVGSKRAMDEYTSIEFMGKNTIVLHNTC 420
DB 361 NVVDMDMNSNAIIVYEGEHPHVVIVKVPYVGSKRAMDEYTSIEFMGKNTIVLHNTC 420
QY 421 EDSLLAAPIIIDLVLAELSTRIOFKANEKGKFSFHPVATIIISLYTAAPLVPECTPVVN 480
DB 421 EDSLLAAPIIIDLVLAELSTRIOFKANEKGKFSFHPVATIIISLYTAAPLVPECTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 10
ID ADS82000 standard; protein, 510 AA.
XX ADS82000;
AC ADS82000;
XX
DB 18-NOV-2004 (first entry)
XX
DE soybean myo-inositol 1-phosphate synthase wild-type 3.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29004JP01.
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
XX
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
FI Hitz WD, Sebastian SA;
```

XX WPI; 2004-639957/62.
DR N-PSDB; ADS81999.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX Example 8; SEQ ID NO 10; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol 1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a wild-type myo-inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 98.6%; Score 2593; DB 8; Length 510;

Best Local Similarity 98.4%; Pred. No. 1.5e-212; Mismatches 4; Indels 0; Gaps 0;

Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 MFENFKVESPNVKKYETETETISVYNYETTELVEHNRNGTQWIVKPSVNYQFKTNTHVP 60
DB 1 MFENFKVECPNVKKTETETISVYNYETTELVEHNRNGTQWIVKPSVKVEFKTNTHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIYANREDISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIYANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
QY 121 IYAPFKSLPWNPDIVFGGMDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
DB 121 IYAPFKSLPWNPDIVFGGMDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
QY 181 YDPDFIANOEERANNVIKGTKEQVQOIIKDIAFKKATKVDKVVLTANTERYSLV 240
DB 181 YDPDFIANOEERANNVIKGTKEQVQOIIKDIAFKKATKVDKVVLTANTERYSLV 240
QY 241 VGLNDTMENLLAAVDRNEAETSPSTLYAIACMVNVPFINGSPONTFVPGILDAIARNT 300
DB 241 VGLNDTMENLLAAVDRNEAETSPSTLYAIACMVNVPFINGSPONTFVPGILDAIARNT 300
QY 301 LIGGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNNGNLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNNGNLSAPQTFRSKEISKS 360
QY 361 NVVDMDMNSNAIIVYEGEHPHVVIVKVPYVGSKRAMDEYTSIEFMGKNTIVLHNTC 420
DB 361 NVVDMDMNSNAIIVYEGEHPHVVIVKVPYVGSKRAMDEYTSIEFMGKNTIVLHNTC 420
```

DB 361 NVDDMVNSNALIYEPGHPDHVVYIKYVPYVGDSKRAMDEYSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPIIIDVLVLAELSTRIOFKAENEGFHSFHPVATILSLTAPLVPPTGTPVNV 480
DB 421 EDSLLAAPIIIDVLVLAELSTRIEFKAENEGFHSFHPVATILSLTAPLVPPTGTPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
RESULT 11
ADS81994
ID ADS81994 standard; protein; 510 AA.
AC ADS81994;
DT 18-NOV-2004 (first entry)
XX Soybean myo-inositol 1-phosphate synthase wild-type 1.
DE
XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX Glycine max; cultivar Wye.
OS
XX US2003074685-A1.
PM 17-APR-2003.
PD 11-MAR-2002; 2002US-00025003.
PF 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
PI WPI; 2004-639957/62.
DR N-PSDB; ADS81993.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
PS Claim 3; SEQ ID NO 2; 34pp; English.

CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a wild-type myo-inositol 1-phosphate synthase.
XX
SQ Sequence 510 AA:
Query Match 98.6%; Score 2593; DB 8; Length 510;
Best Local Similarity 98.4%; Pred. No..1.5e-212;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 MFIENFKYESPNVKKTEIEIOSVYNYETTELVHNRNCTQWYIKPKSVNTOFKNTHTVP 60
DB 1 MFIENFKYECNPVKKTEIEIOSVYNYETTELVHNRNCTQWYIKPKSVKVEFKNTHTVP 60
QY 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVGSFQGE 120
DB 61 KLGWLVGMGNGNSTLTGVIANREGISWATKDKIOQANFGSLTQASAIRVGSFQGE 120
QY 121 IYAPFKSLPMVNPDDIVFGGMDISNMNLADAMARAKVFDIDLOQLRPYMESWVLPGI 180
DB 121 IYAPFKSLPMVNPDDIVFGGMDISNMNLADAMARAKVFDIDLOQLRPYMESWVLPGI 180
QY 181 YDPPFIANOEERANNVKGTQKQOVQOIIDIAFKKATVDKRVVLTANTERTSMLV 240
DB 181 YDPPFIANOEERANNVKGTQKQOVQOIIDIAFKKATVDKRVVLTANTERTSMLV 240
QY 241 VGLNDTMENLAAVDRNEASISPTLYAIACMENVPEPINGSPONTFPGILDLAIANT 300
DB 241 VGLNDTMENLAAVDRNEASISPTLYAIACMENVPEPINGSPONTFPGILDLAIANT 300
QY 301 LIIGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSPQTFRSKEISKS 360
DB 301 LIIGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSPQTFRSKEISKS 360
QY 361 NVDDMVNSNALIYEPGHPDHVVYIKYVPYVGDSKRAMDEYSEIFMGKSTIVLHNTC 420
DB 361 NVDDMVNSNALIYEPGHPDHVVYIKYVPYVGDSKRAMDEYSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPIIIDVLVLAELSTRIOFKAENEGFHSFHPVATILSLTAPLVPPTGTPVNV 480
DB 421 EDSLLAAPIIIDVLVLAELSTRIEFKAENEGFHSFHPVATILSLTAPLVPPTGTPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
RESULT 12
AAW79741
ID AAW79741 standard; protein; 510 AA.
AC AAW79741;
XX 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
XX Soybean mutant myo-inositol 1-phosphate synthase.
KM Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid.
XX Glycine max; line LR33.
OS
PM WO9845448-A1.
XX 15-OCT-1998.
PD 07-APR-1998; 98WO-US006822.
PF 08-APR-1997; 97US-00835751.


```
Db 361 NVVDDMNSNALIYEPGHPDHVVVIKYVPYVGDNSRAMDEYTSIEMGKSTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Db 421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510

RESULT 15
AAZ24477
ID AAZ24477 standard; protein; 536 AA.
AC AAZ24477;
XX
XX 24-SEP-1999 (first entry)
DT
XX Nicotiana paniculata INPS protein.
DE
XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
KW water stress; resistance.
XX Nicotiana paniculata.
OS
XX Key Location/Qualifiers
FH Misc-difference 511
FT /label= unknown
FT /note= "encoded by the stop codon TGA"
FT
XX JP1187879-A.
PN
XX 13-JUL-1999.
PD
XX 26-DEC-1997; 97JP-00359773.
PE
XX 26-DEC-1997; 97JP-00359773.
PR
XX 26-DEC-1997; 97JP-00359773.
XX
XX (NISH ) JAPAN TOBACCO INC.
PA
XX WPI; 1999-451546/38.
DR N-PSDB; AAX90402.
XX
XX New INPS gene derived from Nicotiana genus plant - useful for conferring
PT resistance to water stress to plants.
PS
XX Claim 2; Page 6-8; 8pp; Japanese.
CC The present sequence is the Nicotiana paniculata inositol monophosphate
CC synthase (INPS), designated NpINPS1. INPS can be used to confer water
CC stress resistance to a plant
XX
XX Sequence 536 AA;
SQ

Query Match 93.9%; Score 2470; DB 2; Length 536;
Best Local Similarity 92.4%; Pred. No. 5,4e-202;
Matches 471; Conservative 21; Mismatches 18; Indels 0; Gaps 0;

QY 1 MFIENFKVESPVKKTETETIEISVYNYETTELVHENRNGTYQWIVPKSVNTQFKTNTVHP 60
Db 1 MFIENFKVESPVKKTETETIEISVYNYETTELVHENRNGTYQWIVPKSVNTQFKTNTVHP 60
QY 61 KLGVVLVGMGNGSTLTGAVYANREDISWATKDKIQOANYFGSLTOASAIRVGSFOGEE 120
Db 61 KLGVVLVGMGNGSTLTGAVYANREDISWATKDKIQOANYFGSLTOASAIRVGSFOGEE 120
QY 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
Db 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
QY 181 YDPPFIANOGSRANNVIKGTKEQIDQIILKDIREFEKAKVKVVLMTANTERYSNV 240
Db 181 YDPPFIANOGSRANNVIKGTKEQIDQIILKDIREFEKAKVKVVLMTANTERYSNV 240
```

```
Db 181 YDPPFIANOGSRANNVIKGTKEQIDQIILKDIREFEKAKVKVVLMTANTERYSNV 240
QY 241 VGLNDTMENTLAAVDRNEAEISPTLYAIAQWENVEPINGSPONTFVPGIIDLAIARNT 300
Db 241 VGLNDTMENTLAAVDRNEAEISPTLYAIAQWENVEPINGSPONTFVPGIIDLAIARNT 300
QY 301 LIGDDFKSGQTKKKSIVLDFVGAGIKPTISIVSYNHLGNNDGNLAPQTFRSKEISKS 360
Db 301 LIGDDFKSGQTKKKSIVLDFVGAGIKPTISIVSYNHLGNNDGNLAPQTFRSKEISKS 360
QY 361 NVVDDMNSNALIYEPGHPDHVVVIKYVPYVGDNSRAMDEYTSIEMGKNTIVLHNTC 420
Db 361 NVVDDMNSNALIYEPGHPDHVVVIKYVPYVGDNSRAMDEYTSIEMGKNTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Db 421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
```

Search completed: June 7, 2005, 16:43:37
Job time : 93.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 03:03:45 ; Search time 952.946 Seconds
(without alignments)
9972.402 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 acgtccatcagagattctaa.....tgattctcagracaaatga 1533

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues 12093534

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11A_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1533	100.0	1533	14	US-10-025-003-11
2	1533	100.0	1533	19	US-10-718-952-11
3	1531.4	99.9	1533	14	US-10-025-003-11
4	1531.4	99.9	1533	14	US-10-025-003-11
5	1531.4	99.9	1533	19	US-10-718-952-11
6	1531.4	99.9	1533	19	US-10-718-952-11
7	1464.2	95.5	1533	14	US-10-025-003-9

8	1464.2	95.5	1533	19	US-10-718-952-9	Sequence 9, Appl1
9	1464.2	95.5	1760	14	US-10-025-003-1	Sequence 1, Appl1
10	1464.2	95.5	1760	14	US-10-718-952-1	Sequence 1, Appl1
11	1462.6	95.4	1533	14	US-10-025-003-5	Sequence 5, Appl1
12	1462.6	95.4	1533	19	US-10-718-952-5	Sequence 5, Appl1
13	1456.4	95.0	1989	18	US-10-424-599-70167	Sequence 70167, A
14	1261	82.3	2018	18	US-10-424-599-12022	Sequence 12022, A
15	1176.4	76.7	2582	18	US-10-424-559-12021	Sequence 12021, A
16	1000.2	65.2	1924	19	US-10-767-701-14714	Sequence 14714, A
17	998.6	65.1	1533	9	US-09-938-842A-1438	Sequence 1438, Ap
18	998.6	65.1	1533	11	US-09-938-842A-1438	Sequence 1438, Ap
19	995.4	64.9	1941	20	US-10-739-990-2091	Sequence 2091, Ap
20	990.6	64.6	2284	20	US-10-425-115-47189	Sequence 47189, A
21	989	64.5	2156	18	US-10-425-114-25858	Sequence 25858, A
22	987.4	64.4	2444	20	US-10-425-115-122151	Sequence 122151, A
23	985.8	64.3	1988	18	US-10-425-114-7081	Sequence 7081, Ap
24	984.2	64.2	1879	19	US-10-425-114-5704	Sequence 5704, Ap
25	984.2	64.2	1914	18	US-10-425-114-5612	Sequence 5612, Ap
26	984.2	64.2	1959	9	US-09-727-628-1	Sequence 1, Appl1
27	982.6	64.1	1971	18	US-10-425-114-25983	Sequence 25983, A
28	982.6	64.1	2233	20	US-10-425-115-47187	Sequence 47187, A
29	981	64.0	2309	20	US-10-425-115-47194	Sequence 47194, A
30	981	64.0	2576	19	US-10-437-963-70769	Sequence 70769, A
31	979.4	63.9	1912	18	US-10-425-114-8483	Sequence 8483, Ap
32	979.4	63.9	1931	9	US-09-921-232-10	Sequence 10, Appl1
33	979.4	63.9	1931	9	US-09-921-330-10	Sequence 10, Appl1
34	979.4	63.9	1931	9	US-09-921-329-10	Sequence 10, Appl1
35	979.4	63.9	1997	20	US-10-425-115-47190	Sequence 47190, A
36	970	63.3	2582	18	US-10-424-599-12021	Sequence 12021, A
37	941.6	61.4	1759	19	US-10-363-859-205	Sequence 205, App
38	926.6	60.4	1858	18	US-10-425-114-14009	Sequence 14009, A
39	926.6	60.4	2050	18	US-10-425-114-5864	Sequence 5864, Ap
40	925	60.3	1853	18	US-10-425-114-16524	Sequence 16524, Ap
41	925	60.3	2130	20	US-10-425-115-83553	Sequence 83553, A
42	910.8	59.4	1673	20	US-10-425-115-47180	Sequence 47180, A
43	901	58.8	2196	19	US-10-437-963-85285	Sequence 85285, A
44	891.6	58.2	1789	20	US-10-425-115-47182	Sequence 47182, A
45	665	43.4	1240	18	US-10-425-114-5619	Sequence 5619, Ap

ALIGNMENTS

RESULT 1
US-10-025-003-11
; Sequence 11, Application US/10025003
; Publication No. US2003007465A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streit, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025, 003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835, 751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 11
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-11
Query Match 100.0%; Score 1533; DB 14; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGTCATCGAGAAATTTTAAAGGTAGAGAGTCTTAATGTGAATGATACCGAGACTGAGATT 60
DB 1 ATGTCATCGAGAAATTTTAAAGGTAGAGAGTCTTAATGTGAATGATACCGAGACTGAGATT 60
OY 61 CAGTCGGGTGTAACAATGAGAAACCAACGAACTGTTCACGAGAACAGAAATGGCACTAT 120
DB 61 CAGTCGGGTGTAACAATGAGAAACCAACGAACTGTTCACGAGAACAGAAATGGCACTAT 120
OY 121 CAGTGGATTTGTCAAAACCAAAATCCGCTCAACTACCAATTTTAAACCAACCAACCAATGTCCA 180
DB 121 CAGTGGATTTGTCAAAACCAAAATCCGCTCAACTACCAATTTTAAACCAACCAACCAATGTCCA 180
OY 181 AAATGGGGGTGATGCTTGTGGGTTGGGGTGGAAAACAACGAGCTTACCTCAACCGGTGGT 240
DB 181 AAATGGGGGTGATGCTTGTGGGTTGGGGTGGAAAACAACGAGCTTACCTCAACCGGTGGT 240
OY 241 GTTATTTGCTAACAGAGAGACATTTTCATGAGGCTTACAAAGAGACAAATTCACAACGCCAAT 300
DB 241 GTTATTTGCTAACAGAGAGACATTTTCATGAGGCTTACAAAGAGACAAATTCACAACGCCAAT 300
OY 301 TACTTTGGTCCCTCAACCCAGCTTCAGTATTCAGAGTTGATTCCTTCCAGGAGAGAGA 360
DB 301 TACTTTGGTCCCTCAACCCAGCTTCAGTATTCAGAGTTGATTCCTTCCAGGAGAGAGA 360
OY 361 ATCTATGCCCCATTCAGAGAGTCTGCTTCCATGATTAATCTCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCCATTCAGAGAGTCTGCTTCCATGATTAATCTCTGACGACATTTGTTGGG 420
OY 421 GGATGGGATATGACCAACATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
DB 421 GGATGGGATATGACCAACATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
OY 481 ATCGATTTTCAGAGAGAGTGAAGGCTTTCATGAGATTCATGCTTCCACCTCCCGGAATC 540
DB 481 ATCGATTTTCAGAGAGAGTGAAGGCTTTCATGAGATTCATGCTTCCACCTCCCGGAATC 540
OY 541 TACGACCCCGGATTTCTTGTCTGCAACCAAGAGAGACGTGCAACCAACGTGAATTAAGGCG 600
DB 541 TACGACCCCGGATTTCTTGTCTGCAACCAAGAGAGACGTGCAACCAACGTGAATTAAGGCG 600
OY 601 ACAAAAGCAAGAGCAATGTCAGCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACC 660
DB 601 ACAAAAGCAAGAGCAATGTCAGCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACC 660
OY 661 AAAGTGAACAAGGTGTTGCTCCGTGAGACCTGCCAACAAGAGAGATGCAATTTGGTT 720
DB 661 AAAGTGAACAAGGTGTTGCTCCGTGAGACCTGCCAACAAGAGAGATGCAATTTGGTT 720
OY 721 GTAGGCTTAATGACACATGAGAGATCTTGTGGCTGTGTGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTAATGACACATGAGAGATCTTGTGGCTGTGTGACAGAAATGAGGCTGAG 780
OY 781 ATTTCTCTTCCACCTGTGATGCAATGCTGTGTGATGGAATAATGTTCTTTCATTAAT 840
DB 781 ATTTCTCTTCCACCTGTGATGCAATGCTGTGTGATGGAATAATGTTCTTTCATTAAT 840
OY 841 GGAAGCCCTCAGAACTTTTGTACCAAGGCTGATGATCTTCCATCGGAGAGAACT 900
DB 841 GGAAGCCCTCAGAACTTTTGTACCAAGGCTGATGATCTTCCATCGGAGAGAACT 900
OY 901 TTGATTTGAGAGATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTGAT 960
DB 901 TTGATTTGAGAGATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTGAT 960
OY 961 TTTCTTTGGGGCTGGTATCAAGCCAACTATAGTTAGTTACAAACATCTGGGAAAC 1020
DB 961 TTTCTTTGGGGCTGGTATCAAGCCAACTATAGTTAGTTACAAACATCTGGGAAAC 1020
OY 1021 AATGATGATGATGATCTTGTGGCTCCAAACCTTCCGCTCAAGAGAAATCTCAAGAGC 1080
DB 1021 AATGATGATGATGATCTTGTGGCTCCAAACCTTCCGCTCAAGAGAAATCTCAAGAGC 1080
OY 1081 AAGGTTGTGACATATGTCACAGCAATGCAATCTCTATGAGCTGTGAAACATCCC 1140
```

```
DB 1081 AACGTTGTGACGATATGTCACAGCAATGCAATCCCTATGAGCCGTGGAACATCCC 1140
OY 1141 GACCATGTTGTTATTAAGATATGCTTACGTAAGGAGATAGCAAGAGCCATGAT 1200
DB 1141 GACCATGTTGTTATTAAGATATGCTTACGTAAGGAGATAGCAAGAGCCATGAT 1200
OY 1201 GAGTACACTTGAAGATATTCATGAGTGAAGAAACACCATTTGTTTTCACAAACATGT 1260
DB 1201 GAGTACACTTGAAGATATTCATGAGTGAAGAAACACCATTTGTTTTCACAAACATGT 1260
OY 1261 GAGATTCCTTTTATGCTGCTCTTATTTATCTTGAATCTGTGCTTCTTGTGAGCTGAGC 1320
DB 1261 GAGATTCCTTTTATGCTGCTCTTATTTATCTTGAATCTGTGCTTCTTGTGAGCTGAGC 1320
OY 1321 ACTAATATCCAGTTTAAAGCTGAATAAGGAAATTCACATCTCATTCACCCAGTTGCT 1380
DB 1321 ACTAATATCCAGTTTAAAGCTGAATAAGGAAATTCACATCTCATTCACCCAGTTGCT 1380
OY 1381 ACCATTTCTAGCTATCTGACCAAGGCTCCTGTTTCCACCGGATACCAAGTGTGAT 1440
DB 1381 ACCATTTCTAGCTATCTGACCAAGGCTCCTGTTTCCACCGGATACCAAGTGTGAT 1440
OY 1441 GCATTTGCAAGACGCTGCAATGCTGGAACCAATTAAGAGGCTTGTGTGATTTGGCC 1500
DB 1441 GCATTTGCAAGACGCTGCAATGCTGGAACCAATTAAGAGGCTTGTGTGATTTGGCC 1500
OY 1501 CCAGGAATTAACATGATCTGAGTACAAAGTGA 1533
DB 1501 CCAGGAATTAACATGATCTGAGTACAAAGTGA 1533

RESULT 2
US-10-718-952-11
; Sequence 11, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-718-952-11

Query Match 100.0%; Score 1533; DB 19; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGTCATCGAGAAATTTTAAAGGTAGAGAGTCTTAATGTGAATGATACCGAGACTGAGATT 60
DB 1 ATGTCATCGAGAAATTTTAAAGGTAGAGAGTCTTAATGTGAATGATACCGAGACTGAGATT 60
OY 61 CAGTCGGGTGTAACAATGAGAAACCAACGAACTGTTCACGAGAACAGAAATGGCACTAT 120
DB 61 CAGTCGGGTGTAACAATGAGAAACCAACGAACTGTTCACGAGAACAGAAATGGCACTAT 120
OY 121 CAGTGGATTTGTCAAAACCAAAATCCGCTCAACTACCAATTTTAAACCAACCAACCAATGTCCA 180
DB 121 CAGTGGATTTGTCAAAACCAAAATCCGCTCAACTACCAATTTTAAACCAACCAACCAATGTCCA 180
```

181 AAATTGGGGGTGATGCTTGGGGTTGGGGGGAACAAGGCTCTACCCCTACCGGTGAT 240
181 AAATTGGGGGTGATGCTTGGGGTTGGGGGGAACAAGGCTCTACCCCTACCGGTGAT 240
241 GTTATTGCTTACAGAGAGACATTTTCATGGGCTTACAAAGACATGATTTCAACAGCAAT 300
241 GTTATTGCTTACAGAGAGACATTTTCATGGGCTTACAAAGACATGATTTCAACAGCAAT 300
301 TACTTTGGCTCCCTCACCCCAAGCTCAGCTATTGAGTTGATCTTCCAGGGAGAGAA 360
301 TACTTTGGCTCCCTCACCCCAAGCTCAGCTATTGAGTTGATCTTCCAGGGAGAGAA 360
361 ATCTATGCCCCATTCAAGAGCTGCTTCAATGGTTAACTTCAGACCACTTGTGTTGGG 420
361 ATCTATGCCCCATTCAAGAGCTGCTTCAATGGTTAACTTCAGACCACTTGTGTTGGG 420
421 GGATGGGATATGACAACTGAACCTGGCTGATGCGATGGCCAGGGCAAGGTTGTGAC 480
421 GGATGGGATATGACAACTGAACCTGGCTGATGCGATGGCCAGGGCAAGGTTGTGAC 480
481 ATGATTTGACAGAGAGCTTGAAGGCTTACATGGAATTCATGCTTCCCTCCCGGAATC 540
481 ATGATTTGACAGAGAGCTTGAAGGCTTACATGGAATTCATGCTTCCCTCCCGGAATC 540
541 TACGACCCGGGATTTGATTTGCTGCTCCAAAGAGAGAGCGTCCAAACAGTGTAAAGGC 600
541 TACGACCCGGGATTTGATTTGCTGCTCCAAAGAGAGAGCGTCCAAACAGTGTAAAGGC 600
601 ACAAAGCAAGAGCAAGTTGACGAATCATCAAGACATCAAGCGCTTAAAGAGAGCACC 660
601 ACAAAGCAAGAGCAAGTTGACGAATCATCAAGACATCAAGCGCTTAAAGAGAGCACC 660
661 AAAGTGGACAGAGTGTGCTGCTGTGAGTGCACACAGAGAGGTATGCAATTTGGTT 720
661 AAAGTGGACAGAGTGTGCTGCTGTGAGTGCACACAGAGAGGTATGCAATTTGGTT 720
721 GTAGGCTCTTAATGACACCTGAGAACTCTCTGGCTGCTGTGAGACGAATAGAGCTGAG 780
721 GTAGGCTCTTAATGACACCTGAGAACTCTCTGGCTGCTGTGAGACGAATAGAGCTGAG 780
781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATATTTCTTCTTCAATTAAT 840
781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATATTTCTTCTTCAATTAAT 840
841 GGAAGCCCTCAGAGACCTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAGAACT 900
841 GGAAGCCCTCAGAGACCTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAGAACT 900
901 TTGATTTGGTGAATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTTGGTGAAT 960
901 TTGATTTGGTGAATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTTGGTGAAT 960
961 TTTCTTGGGGGGGTGATTCAGAGCAACATCATATAGTTAACAACCATCTGGGAAAC 1020
961 TTTCTTGGGGGGGTGATTCAGAGCAACATCATATAGTTAACAACCATCTGGGAAAC 1020
1021 AATGATGATGATATCTCGGCTCCAGAACCTTCGCTCCAGAGAAATCTTCAAGAGC 1080
1021 AATGATGATGATATCTCGGCTCCAGAACCTTCGCTCCAGAGAAATCTTCAAGAGC 1080
1081 AAGCTTTGAGAGATGATGATCAACAGCAATGCTCTTATGAGAGCTGTGTAAATCCC 1140
1081 AAGCTTTGAGAGATGATGATCAACAGCAATGCTCTTATGAGAGCTGTGTAAATCCC 1140
1141 GACCATGTTGTTATTATTAAGATGCTTACGTAAGGAGATAGCAAGAGAGCATGAT 1200
1141 GACCATGTTGTTATTATTAAGATGCTTACGTAAGGAGATAGCAAGAGAGCATGAT 1200
1201 GAGTACACTTCAAGAGATATTCATGGGTGAAAGAACACCATTTGTTGCAACACATGT 1260
1201 GAGTACACTTCAAGAGATATTCATGGGTGAAAGAACACCATTTGTTGCAACACATGT 1260

1261 GAGATTCCTTTTATGAGCTGCTCTATATCTTGAAGTGTGCTTCTTGTGAGTGAAC 1320
1261 GAGATTCCTTTTATGAGCTGCTCTATATCTTGAAGTGTGCTTCTTGTGAGTGAAC 1320
1321 ACTAGATTCAGTATTAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
1321 ACTAGATTCAGTATTAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
1381 ACCATTTCTAGCTATCTGACCAAGGCTCCTCTGTGTTCCACCGGGTACACCACTGTGAT 1440
1381 ACCATTTCTAGCTATCTGACCAAGGCTCCTCTGTGTTCCACCGGGTACACCACTGTGAT 1440
1441 GCATTTGCAAG 1500
1441 GCATTTGCAAG 1500
1501 CCAGAGATTAACATGATTTCTGAGTACAGTGA 1533
1501 CCAGAGATTAACATGATTTCTGAGTACAGTGA 1533

RESULT 3
US-10-025-003-13
; Sequence 13, Application us/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAPINOSIDE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-025-003-13

Query Match 99.9%; Score 1531.4; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTAAGTACACCGAGATGAGATT 60
1 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTAAGTACACCGAGATGAGATT 60
61 CAGTCCGTGTACACTACGAAACCAACCACTTGTTCACGAGACAGAGATGCAACTAT 120
61 CAGTCCGTGTACACTACGAAACCAACCACTTGTTCACGAGACAGAGATGCAACTAT 120
61 CAGTCCGTGTACACTACGAAACCAACCACTTGTTCACGAGACAGAGATGCAACTAT 120
121 CAGTGGATTTGCAAAACCAACCACTGCAACTACCAATTTTAAACCAACCACTGTTCA 180
121 CAGTGGATTTGCAAAACCAACCACTGCAACTACCAATTTTAAACCAACCACTGTTCA 180
181 AAATTGGGGGTGATGCTTGTGGTTGGGGTGGAAACAACGCTCTTACCTTCAACCGGTGT 240
181 AAATTGGGGGTGATGCTTGTGGTTGGGGTGGAAACAACGCTCTTACCTTCAACCGGTGT 240
241 GTTATTGCTTACAGAGAGACATTTTCATGGGCTTACAAAGACATGATTTCAACAGCAAT 300
241 GTTATTGCTTACAGAGAGACATTTTCATGGGCTTACAAAGACATGATTTCAACAGCAAT 300
301 TACTTTGGCTCCCTCACCCCAAGCTCAGCTATTGAGTTGATCTTCCAGGGAGAGAA 360
301 TACTTTGGCTCCCTCACCCCAAGCTCAGCTATTGAGTTGATCTTCCAGGGAGAGAA 360

```
Db 301 TACTTGGCTCCCTCAACCCAGCCTCAGCTATTGAGTTGATTCCTTCCAGGAGAGAA 360
Qy ATCTATGCCCATTTCAAGAGTCTGCTTCCATAGTTAATCTGACGACATGTTGTTGGG 420
Db 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATAGTTAATCTGACGACATGTTGTTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
Qy 481 ATCGATTTGACAGAGAGATTGAGGCTTTACATGAGATTCATGTTCCACTCCCGGGAATC 540
Db 481 ATCGATTTGACAGAGAGATTGAGGCTTTACATGAGATTCATGTTCCACTCCCGGGAATC 540
Qy 541 TACGACCCCGGATTTCACTGCTGCAACCAAGAGAGCGTCCCAACACTGATTAAGGCGC 600
Db 541 TACGACCCCGGATTTCACTGCTGCAACCAAGAGAGCGTCCCAACACTGATTAAGGCGC 600
Qy 601 ACAAAGCAAGACAGATTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAAAGCCACC 660
Db 601 ACAAAGCAAGACAGATTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAAAGCCACC 660
Qy 661 AAAGTGGACAAGGTGTTGCTCTGAGACTGCGCAACAGAGAGATATAGCAATTTGGT 720
Db 661 AAAGTGGACAAGGTGTTGCTCTGAGACTGCGCAACAGAGAGATATAGCAATTTGGT 720
Qy 721 GTAGGCTTAATGACACATGAGAAATCTCTGGCTGCTGAGACAGAAATGAGGCTGAG 780
Db 721 GTAGGCTTAATGACACATGAGAAATCTCTGGCTGCTGAGACAGAAATGAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCGCATGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCGCATGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
Qy 841 GGAAGCCCTCAAGAACTTTTGTACAGAGGCTGATGATCTTGCCATCCCGAGAAACAT 900
Db 841 GGAAGCCCTCAAGAACTTTTGTACAGAGGCTGATGATCTTGCCATCCCGAGAAACAT 900
Qy 901 TTGATGTTGAGAGATGACTTCAAGAGTGGTCAAGACCAAAATGAAATCTGTGTTGGAT 960
Db 901 TTGATGTTGAGAGATGACTTCAAGAGTGGTCAAGACCAAAATGAAATCTGTGTTGGAT 960
Qy 961 TTTCTTGTGGGGCTGATCAAGCCAACTATAGTTAAGTTCAACCAATCTGGGAAAC 1020
Db 961 TTTCTTGTGGGGCTGATCAAGCCAACTATAGTTAAGTTCAACCAATCTGGGAAAC 1020
Qy 1021 AATGATGATGAAATCTCTGCGCTCAAAACCTTCCGCTCCAGGAAATCTCCAGAAC 1080
Db 1021 AATGATGATGAAATCTCTGCGCTCAAAACCTTCCGCTCCAGGAAATCTCCAGAAC 1080
Qy 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGGTGAACTCCC 1140
Db 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGGTGAACTCCC 1140
Qy 1141 GACCATGTTGTTGTTAATTAAGTATGTCCTTACGTAAGGGGATAGCAAGAGCCATGAT 1200
Db 1141 GACCATGTTGTTGTTAATTAAGTATGTCCTTACGTAAGGGGATAGCAAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAAGATATTCATGAGTGGTGAAGAAACCAATGTTTGGACCAACATGT 1260
Db 1201 GAGTACACTTCAAGATATTCATGAGTGGTGAAGAAACCAATGTTTGGACCAACATGT 1260
Qy 1261 GAGGATTCCTTTTACTGCTCTATTAATCTTGAAGTGGTCTCTTCTTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTTACTGCTCTATTAATCTTGAAGTGGTCTCTTCTTGAGCTGAGC 1320
Qy 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACACTCATTTCCACCAAGTTGCT 1380
Db 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACACTCATTTCCACCAAGTTGCT 1380
Qy 1381 ACCATTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGCTACCAAGTGGTGAAT 1440
Db 1381 ACCATTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGCTACCAAGTGGTGAAT 1440
```

```
Qy 1441 GCATTGTCAAGCAGCGTGCAATGCTGGAAAAATATAGGCGCTTGTGTTGATTGGCC 1500
Db 1441 GCATTGTCAAGCAGCGTGCAATGCTGGAAAAATATAGGCGCTTGTGTTGATTGGCC 1500
Qy 1501 CCAGAGATTAACATGATTTCTGAGTACAAAGTGA 1533
Db 1501 CCAGAGATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 4
US-10-025-003-15
; Sequence 15, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-15

Query Match 99.9%; Score 1531.4; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTTCAATCGAATTTTAAGTAGAGAGTCTCTAATGTGAATGACCGAGACTGAGATT 60
Db 1 ATGTTCAATCGAATTTTAAGTAGAGAGTCTCTAATGTGAATGACCGAGACTGAGATT 60
Qy 61 CAGTCCGTGTACAACTAGAAACCAACGAACTTGTTCAGAAACAGAAATGGCACTAT 120
Db 61 CAGTCCGTGTACAACTAGAAACCAACGAACTTGTTCAGAAACAGAAATGGCACTAT 120
Qy 121 CAGTGAATTTGAAACCCAAATCCGTCAACTACCAATTTAAACCAACCACTGTTCA 180
Db 121 CAGTGAATTTGAAACCCAAATCCGTCAACTACCAATTTAAACCAACCACTGTTCA 180
Qy 181 AAATTTGGGGGTGATGCTTGTGGGTGGGGTGAACCAAGGCTTACCTCAACCGGTGAT 240
Db 181 AAATTTGGGGGTGATGCTTGTGGGTGGGGTGAACCAAGGCTTACCTCAACCGGTGAT 240
Qy 241 GTTATTTGTAACAGAGAGACATTTCAATGAGGCTACAAAGACAAAGATTCAACAGCAAT 300
Db 241 GTTATTTGTAACAGAGAGACATTTCAATGAGGCTACAAAGACAAAGATTCAACAGCAAT 300
Qy 301 TACTTGGCTCCCTCAACCAAGCTTCAAGCTATTCGAGTTGGATCTTCCAGGAGAGAA 360
Db 301 TACTTGGCTCCCTCAACCAAGCTTCAAGCTATTCGAGTTGGATCTTCCAGGAGAGAA 360
Qy 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAAGTTAATCTGAGCAATGTGTTGGG 420
Db 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAAGTTAATCTGAGCAATGTGTTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
Qy 481 ATCGATTTGACAGAGAGATTGAGGCTTTACATGAGATTCATGTTCCACTCCCGGGAATC 540
```


Db 481 ATGATTTGAGAGAGGTTGAGGCTTATCATGGATTCATGTTCCACTCCCGGAATC 540
Qy 541 TAGACCCCGGATTTCTTGTCTGCCAACCAAGAGAGGCTGCCAACACGTGATTAAGGCG 600
Db 541 TAGACCCCGGATTTCTTGTCTGCCAACCAAGAGAGGCTGCCAACACGTGATTAAGGCG 600
Qy 601 ACAAGCAAGAGAGGTTGAGCAATCATCAAGAGACATCAAGGCGTTTAAAGAGGACCC 660
Db 601 ACAAGCAAGAGAGGTTGAGCAATCATCAAGAGACATCAAGGCGTTTAAAGAGGACCC 660
Qy 661 AAAGTGAGAGAGGTTGTTGCTCTGTGAGTCCCAACAGAGAGGATATAGCAATTTGGTT 720
Db 661 AAAGTGAGAGAGGTTGTTGCTCTGTGAGTCCCAACAGAGAGGATATAGCAATTTGGTT 720
Qy 721 GTAGGCTTAAATGACACATGAGAAATCTTGTGCTGTGAGACAGAAATGAGGCTGAG 780
Db 721 GTAGGCTTAAATGACACATGAGAAATCTTGTGCTGTGAGACAGAAATGAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTTATGACATTTGCTGTGATGAGAAATGCTTCTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTTATGACATTTGCTGTGATGAGAAATGCTTCTTCAATTAAT 840
Qy 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATTTGATCTTGCATCCGAGAGAACT 900
Db 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATTTGATCTTGCATCCGAGAGAACT 900
Qy 901 TTGATTTGAGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 960
Db 901 TTGATTTGAGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 960
Qy 961 TTTCTTGTGGGGCTGTATCAAGGCAATCTATAGTTAGTTACCAACCTCTGGGAAC 1020
Db 961 TTTCTTGTGGGGCTGTATCAAGGCAATCTATAGTTAGTTACCAACCTCTGGGAAC 1020
Qy 1021 AATGATGATGATGATCTCTGAGTCCCAAAACCTTCCGCTCCAAAGAAATCTCCAGAG 1080
Db 1021 AATGATGATGATGATCTCTGAGTCCCAAAACCTTCCGCTCCAAAGAAATCTCCAGAG 1080
Qy 1081 AAGCTTTGAGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1140
Db 1081 AAGCTTTGAGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1140
Qy 1141 GACCATGTTGTTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTT 1200
Db 1141 GACCATGTTGTTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTT 1200
Qy 1201 GAGTACATTTGAGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1260
Db 1201 GAGTACATTTGAGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1260
Qy 1261 GAGGATTTCCCTTTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTT 1320
Db 1261 GAGGATTTCCCTTTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTT 1320
Qy 1321 ACTAGATTCAGTTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTT 1380
Db 1321 ACTAGATTCAGTTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTTAAAGTATGCTT 1380
Qy 1381 ACCATTTCTGAGTATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1440
Db 1381 ACCATTTCTGAGTATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1440
Qy 1441 GCATTTCTCAAGAGTATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1500
Db 1441 GCATTTCTCAAGAGTATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1500
Qy 1501 CGAGAGAAATGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1533
Db 1501 CGAGAGAAATGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 1533

US-10-718-952-13
; Sequence 13, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-13

Query Match 99.9%; Score 1531.4; DB 19; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AGTTCATGAGAGATTTTAAAGTATGAGAGTCTTAAATGTAAGTACCGAGACTGAGATT 60
Db 1 AGTTCATGAGAGATTTTAAAGTATGAGAGTCTTAAATGTAAGTACCGAGACTGAGATT 60
Qy 61 CAGTCCGTGATACATACAGAAACCAACGAACTTGTTCAGAGAACAGAAATGACACTAT 120
Db 61 CAGTCCGTGATACATACAGAAACCAACGAACTTGTTCAGAGAACAGAAATGACACTAT 120
Qy 121 CAGTGGATTTGATCAACCCAAATCCGTCACTACCAATTTTAAACCAACCCATGTTCCA 180
Db 121 CAGTGGATTTGATCAACCCAAATCCGTCACTACCAATTTTAAACCAACCCATGTTCCA 180
Qy 181 AAATGGGGGTATGATGTTGGGGTGGGGTGAACCAAGGCTTACCTCAACCGGTGAT 240
Db 181 AAATGGGGGTATGATGTTGGGGTGGGGTGAACCAAGGCTTACCTCAACCGGTGAT 240
Qy 241 GTTATTTCTAACAGAGAGATTTTATGAGGCTTACCAAGAGAGATTTTCAAGAGCAAT 300
Db 241 GTTATTTCTAACAGAGAGATTTTATGAGGCTTACCAAGAGAGATTTTCAAGAGCAAT 300
Qy 301 TACTTTGGCTCCCTCAACCCAGAGGCTTATTCAGAGTGTGATCTTCCAGAGAGAGAA 360
Db 301 TACTTTGGCTCCCTCAACCCAGAGGCTTATTCAGAGTGTGATCTTCCAGAGAGAGAA 360
Qy 361 ACTATGAGGATGATGAGAGTGTGCTTCAAGAGTGTGATCTTCCAGAGAGATTTGAGG 420
Db 361 ACTATGAGGATGATGAGAGTGTGCTTCAAGAGTGTGATCTTCCAGAGAGATTTGAGG 420
Qy 421 GGATGGATATGAGAGATGAGAGTGTGCTTCAAGAGTGTGATCTTCCAGAGAGATTTGAG 480
Db 421 GGATGGATATGAGAGATGAGAGTGTGCTTCAAGAGTGTGATCTTCCAGAGAGATTTGAG 480
Qy 481 ATGATTTGAGAGAGATGAGAGTGTGCTTCAAGAGTGTGATCTTCCAGAGAGATTTGAG 540
Db 481 ATGATTTGAGAGAGATGAGAGTGTGCTTCAAGAGTGTGATCTTCCAGAGAGATTTGAG 540
Qy 541 TAGAGCCGGATTTTATGCTGCAACCAAGAGAGGCGTGAACCAAGAGATTTTAAAGGCG 600
Db 541 TAGAGCCGGATTTTATGCTGCAACCAAGAGAGGCGTGAACCAAGAGATTTTAAAGGCG 600
Qy 601 ACAAGCAAGAGAGATTTTATGCTGCAACCAAGAGAGGCGTGAACCAAGAGATTTTAAAGGCG 660
Db 601 ACAAGCAAGAGAGATTTTATGCTGCAACCAAGAGAGGCGTGAACCAAGAGATTTTAAAGGCG 660

QY 661 AAAGTGACAAGGTGGTGTCTGCTGAGACTGCGCAACACAGAGAGGTATAGCAATTTGGT 720
DB 661 AAAGTGACAAGGTGGTGTCTGCTGAGACTGCGCAACACAGAGAGGTATAGCAATTTGGT 720
QY 721 GTAGGCTTTAATGACACATGAGAAATCTCTTGGCTGCTGTGAGCAAGAAATAGAGCTGAG 780
DB 721 GTAGGCTTTAATGACACATGAGAAATCTCTTGGCTGCTGTGAGCAAGAAATAGAGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGATGACCATGCTGCTGAGTGAAGAAATGTTCTTCTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTGATGACCATGCTGCTGAGTGAAGAAATGTTCTTCTTCAATTAAT 840
QY 841 GGAAGCCCTCAGAGACATTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAACT 900
DB 841 GGAAGCCCTCAGAGACATTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAACT 900
QY 901 TTGATTTGGTGAAGTACTTCAAGAGTGTCAAGCCAAATGAATCTGTGTTGGTGAAT 960
DB 901 TTGATTTGGTGAAGTACTTCAAGAGTGTCAAGCCAAATGAATCTGTGTTGGTGAAT 960
QY 961 TTTCTTGGGGGGGCTGATCAAGCCACATCTATGATTAACAACCATCGGGAAAC 1020
DB 961 TTTCTTGGGGGGGCTGATCAAGCCACATCTATGATTAACAACCATCGGGAAAC 1020
QY 1021 AATGATGATGATATCTCTCGGCTCCACAAACCTTCCGCTCAGAGAAATCTCAAGAC 1080
DB 1021 AATGATGATGATATCTCTCGGCTCCACAAACCTTCCGCTCAGAGAAATCTCAAGAC 1080
QY 1081 AAGCTTTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAACATCCC 1140
DB 1081 AAGCTTTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAACATCCC 1140
QY 1141 GACCAATGTTGTTATTAAGATATGACCTTAAGTGGGGAATAGCAAGAGGCAATGAT 1200
DB 1141 GACCAATGTTGTTATTAAGATATGACCTTAAGTGGGGAATAGCAAGAGGCAATGAT 1200
QY 1201 GAGTACACTTCAAGATATTCATGAGGTGAGAAAGAACCAATGTTTGGCAACAACATGT 1260
DB 1201 GAGTACACTTCAAGATATTCATGAGGTGAGAAAGAACCAATGTTTGGCAACAACATGT 1260
QY 1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGGACTTGGTCTTGTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGGACTTGGTCTTGTGAGCTGAGC 1320
QY 1321 ACTAGAAATCAGTTTAAAGCTGAAATGAGGGAATTCACATCTTCCACCCAGTTGCT 1380
DB 1321 ACTAGAAATCAGTTTAAAGCTGAAATGAGGGAATTCACATCTTCCACCCAGTTGCT 1380
QY 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAAT 1440
DB 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAAT 1440
QY 1441 GCATTGTCAAGAGAGGCTGCAATGCTGGAAGAAATGAAGGCTTGTGTTGATTTGGCC 1500
DB 1441 GCATTGTCAAGAGAGGCTGCAATGCTGGAAGAAATGAAGGCTTGTGTTGATTTGGCC 1500
QY 1501 CCAGAGAAATPACTGATCTCGATCAAGTGA 1533
DB 1501 CCAGAGAAATPACTGATCTCGATCAAGTGA 1533

RESULT 6

US-10-718-952-15
; Sequence 15, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C

; CURRENT APPLICATION NUMBER: US/10/718, 952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
us-10-718-952-15

Query Match 99.9%; Score 1531.4; DB 19; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTCAATCAGAAATTTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
DB 1 ATGTTCAATCAGAAATTTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
QY 61 CAGTCCGCTGTAACACTAGAAACCAACCGAATCTGTTCAAGAAAGGAACTGCTAT 120
DB 61 CAGTCCGCTGTAACACTAGAAACCAACCGAATCTGTTCAAGAAAGGAACTGCTAT 120
QY 121 CAGTGAATTTGCAACCCAAATCCGTCACATCAATTTAAACCAACCAACCCATGTTCA 180
DB 121 CAGTGAATTTGCAACCCAAATCCGTCACATCAATTTAAACCAACCAACCCATGTTCA 180
QY 181 AAATGGGGGTGATCTGTGGGTTGGGGTGAACCAACCGCTTACCTTACCGGTGT 240
DB 181 AAATGGGGGTGATCTGTGGGTTGGGGTGAACCAACCGCTTACCTTACCGGTGT 240
QY 241 GTTATTCCTAAGAGAGAGACATTTATGAGGCTTACAAAGGAACAAGATTTCAAGCAAT 300
DB 241 GTTATTCCTAAGAGAGAGACATTTATGAGGCTTACAAAGGAACAAGATTTCAAGCAAT 300
QY 301 TACTTGGCTCCCTCACCACCAAGCTCAGCTATTCGAGTTGATTCCTTCCAGGAGAGAA 360
DB 301 TACTTGGCTCCCTCACCACCAAGCTCAGCTATTCGAGTTGATTCCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCCATTCAAGAGTCTGCTTCCATGTTTAATCTGACGACATTTGTGTTGG 420
DB 361 ATCTATGCCCCATTCAAGAGTCTGCTTCCATGTTTAATCTGACGACATTTGTGTTGG 420
QY 421 GGAATGGATATGAGCAATGAACTGCTGATGTCATGCGCAAGGCAAGAGGTTGAC 480
DB 421 GGAATGGATATGAGCAATGAACTGCTGATGTCATGCGCAAGGCAAGAGGTTGAC 480
QY 481 ATCGATTTGCAAGAGAGTTGAGGCTTACATGGAATCCATGTTCCACTCCCGGAATC 540
DB 481 ATCGATTTGCAAGAGAGTTGAGGCTTACATGGAATCCATGTTCCACTCCCGGAATC 540
QY 541 TACGACCCGGAATTTCAATGCTGCTGCAACCAAGAGAGCTGCCAACAAGTATTAAGGCG 600
DB 541 TACGACCCGGAATTTCAATGCTGCTGCAACCAAGAGAGCTGCCAACAAGTATTAAGGCG 600
QY 601 ACAAGCAAGAGAGAGTTGAGCAATCAATCAAGAGCATCAAGCGCTTTAAGAGAGCACC 660
DB 601 ACAAGCAAGAGAGAGTTGAGCAATCAATCAAGAGCATCAAGCGCTTTAAGAGAGCACC 660
QY 661 AAAGTGACAAGGTGGTGTCTGCTGAGACTGCGCAACACAGAGAGGTATAGCAATTTGGTT 720
DB 661 AAAGTGACAAGGTGGTGTCTGCTGAGACTGCGCAACACAGAGAGGTATAGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACATGAGAAATCTTGGCTGCTGTGAGCAAGAAATAGAGCTGAG 780
DB 721 GTAGGCTTTAATGACACATGAGAAATCTTGGCTGCTGTGAGCAAGAAATAGAGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGATGACCATGCTGCTGAGTGAAGAAATGTTCTTCTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTGATGACCATGCTGCTGAGTGAAGAAATGTTCTTCTTCAATTAAT 840

```

QY      841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATCTTGGCATCCGAGGAACACT 900
      |||
Db      841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATCTTGGCATCCGAGGAACACT 900
QY      901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGTTGAT 960
      |||
Db      901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGTTGAT 960
QY      961 TTTCTTTGGGGGCTGTGATCAAGCCAACTATAGTTAGTTACCAACATCTGGGAAAC 1020
      |||
Db      961 TTTCTTTGGGGGCTGTGATCAAGCCAACTATAGTTAGTTAGTTACCAACATCTGGGAAAC 1020
QY      1021 AATGATGATGATATCTCTCGGCTCCAAACCTTCCGCTCCAGAGAAATCTTCCAGAGC 1080
      |||
Db      1021 AATGATGATGATATCTCTCGGCTCCAAACCTTCCGCTCCAGAGAAATCTTCCAGAGC 1080
QY      1081 AACGTTGTGACGATATGATCAACAGCAATGCCATCCTCTATGAGGCTGGTGAACATCC 1140
      |||
Db      1081 AACGTTGTGACGATATGATCAACAGCAATGCCATCCTCTATGAGGCTGGTGAACATCC 1140
QY      1141 GACCATGTTGTTGTTATTAAGTATGTCTTACGTAAGGAGATGACAAAGAGCCATGAT 1200
      |||
Db      1141 GACCATGTTGTTGTTATTAAGTATGTCTTACGTAAGGAGATGACAAAGAGCCATGAT 1200
QY      1201 GAGTACACTTCAGAGATATTCATGGGTGAGAAAGAACATGTTTTCACAAACATATG 1260
      |||
Db      1201 GAGTACACTTCAGAGATATTCATGGGTGAGAAAGAACATGTTTTCACAAACATATG 1260
QY      1261 GAGGATTCCTTTTAAAGCTCTCTATTAATCTTGAATGCTTCTTCTGAGCTGAGC 1320
      |||
Db      1261 GAGGATTCCTTTTAAAGCTCTCTATTAATCTTGAATGCTTCTTCTGAGCTGAGC 1320
QY      1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATTCACACCAAGTTGCT 1380
      |||
Db      1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATTCACACCAAGTTGCT 1380
QY      1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGCTTCCACCGGATCACCAAGTGTGAT 1440
      |||
Db      1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGCTTCCACCGGATCACCAAGTGTGAT 1440
QY      1441 GCATTGTCAAGCAGCGTCAATGCTGAAACATATGAGGCTTGTGTGATTTGGCC 1500
      |||
Db      1441 GCATTGTCAAGCAGCGTCAATGCTGAAACATATGAGGCTTGTGTGATTTGGCC 1500
QY      1501 CCAGAGATTAACATGATTTCTCGAGTACAAATGA 1533
      |||
Db      1501 CCAGAGATTAACATGATTTCTCGAGTACAAATGA 1533

```

RESULT 7

```

US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; PRIOR APPLICATION NUMBER: 2002-05-07
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1533
; TYPE: DNA

```

ORGANISM: Glycine max
US-10-025-003-9

Query Match 95.5%; Score 1464.2; DB 14; Length 1533;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

QY      1 ATGTTTATGAGAAATTTTAAAGTAAAGTCTCTAATGCAATGACCCGAGCTGAGAT 60
      |||
Db      1 ATGTTTATGAGAAATTTTAAAGTAAAGTCTCTAATGCAATGACCCGAGCTGAGAT 60
QY      61 CAGTCCGTGACAACTACGAAACCAACCGAATCTGTTTACAGAAACAGAAATGACACTAT 120
      |||
Db      61 CAGTCCGTGACAACTACGAAACCAACCGAATCTGTTTACAGAAACAGAAATGACACTAT 120
QY      121 CAGTGGATTTGCAAAACCAAAATCCGTCAATCAATTTTAAACCAACCAATGTTTCA 180
      |||
Db      121 CAGTGGATTTGCAAAACCAAAATCCGTCAATCAATTTTAAACCAACCAATGTTTCA 180
QY      181 AATTTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTACCCGTGGT 240
      |||
Db      181 AATTTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTACCCGTGGT 240
QY      241 GTTATTTGCTAACAGAGAGACATTTTCAATGAGCTAACAAAGACAAAGATTCACAGCCAT 300
      |||
Db      241 GTTATTTGCTAACAGAGAGACATTTTCAATGAGCTAACAAAGACAAAGATTCACAGCCAT 300
QY      301 TACTTTGGCTCCCTACCAACCAAGCTCAAGCTATCCAGTTGGGCTCTTCCAGGAGAGAA 360
      |||
Db      301 TACTTTGGCTCCCTACCAACCAAGCTCAAGCTATCCAGTTGGGCTCTTCCAGGAGAGAA 360
QY      361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAATGTTAATCTTGAAGACATTTGTGTTGG 420
      |||
Db      361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAATGTTAATCTTGAAGACATTTGTGTTGG 420
QY      421 GATGGGATATACAGCAACATGAACCTGCTGATGTCATGCGCAGGAGCAAAAGGTTTGA 480
      |||
Db      421 GATGGGATATACAGCAACATGAACCTGCTGATGTCATGCGCAGGAGCAAAAGGTTTGA 480
QY      481 ATCGATTTGCAAGACAGTGAAGGCTTACATGAATTCATGATTCCTTCCATCCCGGAATC 540
      |||
Db      481 ATCGATTTGCAAGACAGTGAAGGCTTACATGAATTCATGATTCCTTCCATCCCGGAATC 540
QY      541 TACGACCGGATTTTCAATGCTGCTCCAAACAGAGAGCGTGCACAACTGATTAAGGCG 600
      |||
Db      541 TACGACCGGATTTTCAATGCTGCTCCAAACAGAGAGCGTGCACAACTGATTAAGGCG 600
QY      601 ACAAGCAAGAGCAAGTTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAGCACCC 660
      |||
Db      601 ACAAGCAAGAGCAAGTTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAGCACCC 660
QY      661 AAAGTGAACAAGGTGTTCTCTGTGACTGCAACACAGAGAGTATGCAATTTGTT 720
      |||
Db      661 AAAGTGAACAAGGTGTTCTCTGTGACTGCAACACAGAGAGTATGCAATTTGTT 720
QY      721 GTAGGCTTAAATGACCAATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
      |||
Db      721 GTAGGCTTAAATGACCAATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY      781 AATTTCCCTTCCACCTTGATGACATGCTGCTGTGAGGAAATGTTTCTTCAATTAAT 840
      |||
Db      781 AATTTCCCTTCCACCTTGATGACATGCTGCTGTGAGGAAATGTTTCTTCAATTAAT 840
QY      841 GGAAGCCCTCAGAACATTTTGTACCAAGGCTGATTTGATCTTGGCATCCGAGGAACACT 900
      |||
Db      841 GGAAGCCCTCAGAACATTTTGTACCAAGGCTGATTTGATCTTGGCATCCGAGGAACACT 900
QY      901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGTTGAT 960
      |||
Db      901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGTTGAT 960
QY      961 TTTCTTTGGGGGCTGTGATCAAGCCAACTATAGTTAGTTAGTTACCAACATCTGGGAAAC 1020
      |||

```

Db 961 TTCCTTGTGGGGGCTGTATCAAGCCAACTATAGTCAGTTACCAACCATCTGGGAAAC 1020
Qy 1021 AATGATGATGAATCTCTGGGCTCCAGAACTTCGCTCCAGGAAATCTCCAAGAGC 1080
Db 1021 AATGATGATGAATCTCTGGGCTCCAGAACTTCGCTCCAGGAAATCTCCAAGAGC 1080
Qy 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCC 1140
Db 1081 AACGTTGTGATGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCA 1140
Qy 1141 GACCATGTTGTGTTATTAAGATGTGCTTACGTAAGGGGATACGAAGAGCCATGAT 1200
Db 1141 GACCATGTTGTGTTATTAAGATGTGCTTACGTAAGGGGATACGAAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAACACCATGTTTGTGCAACAACATGT 1260
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAACACCATGTTTGTGCAACAACATGC 1260
Qy 1261 GAGATTCCTCTTAAAGCTGCTCTATTAATCTTGAATGATGATCTTGTGCTGAGCTGAC 1320
Db 1261 GAGATTCCTCTTAAAGCTGCTCTATTAATCTTGAATGATGATGATGATGATGATGAT 1320
Qy 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATCTCATTCACCAAGTTGCT 1380
Db 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATCTCATTCACCAAGTTGCT 1380
Qy 1381 ACCATTCAGCATATGACCAAGGCTCTGTTGTCACCGGGTACACAGATGATGAT 1440
Db 1381 ACCATTCAGCATATGACCAAGGCTCTGTTGTCACCGGGTACACAGATGATGATGAT 1440
Qy 1441 GCATTGTCAAGCAGCTGCAATGCTGAAACATTAATGAGGCTTGTGTGATTTGGCC 1500
Db 1441 GCATTGTCAAGCAGCTGCAATGCTGAAACATTAATGAGGCTTGTGTGATTTGGCC 1500
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1533
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1533

RESULT 8

US-10-718-952-9
Sequence 9, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAPEINASE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-718-952-9

Query Match 95.5%; Score 1464.2; DB 19; Length 1533;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 ATGTTATGAGAAATTTTAAAGTAGAGTCTTAATGTGAAGTACCGAGACTGAGATT 60
Db 1 ATGTTATGAGAAATTTTAAAGTAGAGTCTTAATGTGAAGTACCGAGACTGAGATT 60

Qy 61 CAGTCCGATACAACTAGAAAACACCGAATCTTTCAGAGAAAGGAATGGCACTAT 120
Db 61 CAGTCCGATACAACTAGAAAACACCGAATCTTTCAGAGAAAGGAATGGCACTAT 120
Qy 121 CAGTGGATTTGCAAAACCAATCCGTCAACTACAAATTTTAAACCAACACCAATGTTCCA 180
Db 121 CAGTGGATTTGCAAAACCAATCCGTCAACTACAAATTTTAAACCAACATCATGTTCTT 180
Qy 181 AAATGGGGGATGATCTTGTGGGTTGGGGTGAACCAACGCGCTACCCCTGACCGGTGT 240
Db 181 AAATGGGGGATGATCTTGTGGGTTGGGGTGAACCAACGCGCTACCCCTGACCGGTGT 240
Qy 241 GTTATGTCTAAACAGAGAGCAATTTCAATGGGCTACAAAGGACAGATTCAACAGCAAT 300
Db 241 GTTATGTCTAAACAGAGAGGCAATTTCAATGGGCTACAAAGGACAGATTCAACAGCAAT 300
Qy 301 TACTTTGGCTCCCTACCCCAAGCCCTGACCTATTCGAGTTGGAATCTTCCAGGAGAGGA 360
Db 301 TACTTTGGCTCCCTACCCCAAGCCCTGACCTATTCGAGTTGGAATCTTCCAGGAGAGGA 360
Qy 361 ATCTATGCCCATTTCAAGAGTCTGCTCAATGTTAAATCCTGACGACATTTGTTGGG 420
Db 361 ATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTAAATCCTGACGACATTTGTTGGG 420
Qy 421 GGATGGATATCAGCAACATGAACCTGGCTGATGCTATGCTGAGGCAAGGCTTTGAC 480
Db 421 GGATGGATATCAGCAACATGAACCTGGCTGATGCTATGCTGAGGCAAGGCTTTGAC 480
Qy 481 ATCATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCCGGGAATC 540
Db 481 ATCATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCCGGGAATC 540
Qy 541 TAGACCCGGATTTTCAATGCTGCTCAACCAAGAGAGGCTGCAACAGTGTAAAGGAC 600
Db 541 TAGACCCGGATTTTCAATGCTGCTCAACCAAGAGAGGCTGCAACAGTGTAAAGGAC 600
Qy 601 ACAAAGCAAGAGCAATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 660
Db 601 ACAAAGCAAGAGCAATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 660
Qy 661 AAAGTGACAAAGTGGTGTCTGTGACCTGCTGCAACAGAGGTTATGCAATTTGGT 720
Db 661 AAAGTGACAAAGTGGTGTCTGTGACCTGCTGCAACAGAGGTTATGCAATTTGGT 720
Qy 721 GTAGGCTTAAATGACCAATGAGAAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 GTAGGCTTAAATGACCAATGAGAAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAAATGTTCTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAAATGTTCTTCAATTAAT 840
Qy 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATTTGATCTTGCATGCGGAGAACCT 900
Db 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATTTGATCTTGCATGCGGAGAACCT 900
Qy 901 TTGATTTGGAGATGACTTCAAGAGTGTGACCAACCAATGAAATCTGTGTTGAT 960
Db 901 TTGATTTGGAGATGACTTCAAGAGTGTGACCAACCAATGAAATCTGTGTTGAT 960
Qy 961 TTTCTTGTGGGGGCTGTATCAAGCAACATCTATAGTTAGTTACCAACCATCTGGGAAAC 1020
Db 961 TTTCTTGTGGGGGCTGTATCAAGCAACATCTATAGTTAGTTACCAACCATCTGGGAAAC 1020
Qy 1021 AATGATGATGAATCTCTGGGCTCCAGAACTTCGCTCCAGGAAATCTCCAAGAGC 1080
Db 1021 AATGATGATGAATCTCTGGGCTCCAGAACTTCGCTCCAGGAAATCTCCAAGAGC 1080
Qy 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCC 1140
Db 1081 AACGTTGTGATGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCA 1140
Qy 1141 GACCATGTTGTGTTATTAAGTATGTGCTTACGTAAGGGGATAGCAAGAGCCATGAT 1200

Db 1141 GACCACTGTTGTTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAGAGATTAATTCATGAGGTGGAAGAACACATGTTTTCACACATGT 1260
Db 1201 GAGTACACTTCAGAGATTAATTCATGAGGTGGAAGAACACATGTTTTCACACATGT 1260
Qy 1261 GAGGATTCCTTTTACTGCTCTTATTAATCTTGAGACTTGGTCTTCTTGTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTTACTGCTCTTATTAATCTTGAGACTTGGTCTTCTTGTGAGCTGAGC 1320
Qy 1321 ACTAGATTCAGTTTAAAGCTGAATAATGAGGAAAAATTCACATCTTCCACCGATGGCT 1380
Db 1321 ACTAGATTCAGTTTAAAGCTGAATAATGAGGAAAAATTCACATCTTCCACCGATGGCT 1380
Qy 1381 ACCATTCCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAAT 1440
Db 1381 ACCATTCCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAAT 1440
Qy 1441 GCATTGTCAAAGAGCGGTGCAATGCTGGAATAATATGAGGCTTGTGTGATGGCC 1500
Db 1441 GCATTGTCAAAGAGCGGTGCAATGCTGGAATAATATGAGGCTTGTGTGATGGCC 1500
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACAAGTGA 1533
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACAAGTGA 1533

RESULT 9

US-10-025-003-1
; Sequence 1, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE OF INVENTION: SACHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-1

Query Match 95.5%; Score 1464.2; DB 14; Length 1760;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
Qy 1 ATGTTATCGAGAAATTTAAGTAGAGAGTCTTAATGTGAATGACACCGAGCTGAGATT 60
Db 54 ATGTTATCGAGAAATTTAAGTAGAGTCTTAATGTGAATGACACCGAGCTGAGATT 113
Qy 61 CAGTCCGTGTACAACTAGCAAAACACCGAATTGTTACGAGAAACAGAAATGGCACTTAT 120
Db 114 CAGTCCGTGTACAACTAGCAAAACACCGAATTGTTACGAGAAACAGAAATGGCACTTAT 173
Qy 121 CAGTGTATGTGAACCAACCAATCCGTCACTACCAATTTAAACCAACCCATGTTTCA 180
Db 174 CAGTGTATGTGAACCAACCAATCCGTCACTACCAATTTAAACCAACCAATGTTTCT 233
Qy 181 AATTTGGGGGTATGCTTGTGGGTGGGGTGAACCAAGGCTCTACCCGTGAGT 240
Db 234 AATTTGGGGGTATGCTTGTGGGTGGGGTGAACCAAGGCTCTACCCGTGAGT 293

Qy 241 GTTATTTGCTAACAGAGAGCAATTCATAGGCTTACAAAGAGCAATTCACAGCCAT 300
Db 294 GTTATTTGCTAACAGAGAGCAATTCATAGGCTTACAAAGAGCAATTCACAGCCAT 353
Qy 301 TACTTTGGCTCCCTCAACCAACCTCAGCTATTCAGATTGATTCCTTCCAGGAGAGAA 360
Db 354 TACTTTGGCTCCCTCAACCAACCTCAGCTATTCAGATTGATTCCTTCCAGGAGAGAA 413
Qy 361 ATCTATGCCCATTTCAAGAGTGTCTTCCATATGTTAATCTGACGACATTTGTTGGG 420
Db 414 ATCTATGCCCATTTCAAGAGTGTCTTCCATATGTTAATCTGACGACATTTGTTGGG 473
Qy 421 GGATGGATATATGACCAATGAACCTGCTGATGCAATGCGCAAGGCAAAAGTGTGAC 480
Db 474 GGATGGATATATGACCAATGAACCTGCTGATGCAATGCGCAAGGCAAAAGTGTGAC 533
Qy 481 ATCGATTTGCAAGAGAGATGAGGCTTACATGGAATCAGATGTTTCACTCCCGGAATC 540
Db 534 ATCGATTTGCAAGAGAGATGAGGCTTACATGGAATCAGATGTTTCACTCCCGGAATC 593
Qy 541 TACGACCCGGAATTTGATGCTGCCAACCAAGAGAGCGTGCCAAACGTTGATTAAGGCG 600
Db 594 TATGACCCGGAATTTGATGCTGCCAACCAAGAGAGCGTGCCAAACGTTGATTAAGGCG 653
Qy 601 ACMAAGCAAGACAAAGTTGACGAATCATCAAAAGCATAGAGCGTTTAAAGAGCGACC 660
Db 654 ACMAAGCAAGACAAAGTTGACGAATCATCAAAAGCATAGAGCGTTTAAAGAGCGACC 713
Qy 661 AAGTGGACAAAGTGTGTCTGAGACGTGCCAACAGAGAGGATAGCAATTTGGTT 720
Db 714 AAGTGGACAAAGTGTGTGTGAGACGTGCCAACAGAGAGGATAGCAATTTGGTT 773
Qy 721 GTAGGCTTTAATGACACATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGCGTGA 780
Db 774 GTAGGCTTTAATGACACATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGCGTGA 833
Qy 781 ATTTTCTCTTCCACTTGTATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATAT 840
Db 834 ATTTTCTCTTCCACTTGTATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATAT 893
Qy 841 GGAAGCCCTCAGAACACTTTGTACAGAGGCGATTTGATCTTGCATGCGAGGAACAT 900
Db 894 GGAAGCCCTCAGAACACTTTGTACAGAGGCGATTTGATCTTGCATGCGAGGAACAT 953
Qy 901 TTGATTTGGTGAATGACTTCAAGAGTGTGACCAACCAATGAATCTGTGTTGAT 960
Db 954 TTGATTTGGTGAATGACTTCAAGAGTGTGACCAACCAATGAATCTGTGTTGAT 1013
Qy 961 TTTCTTGTGGGGGCTGTATACAGCCAACTATATGTTAGTTACAACTGTGGAAAC 1020
Db 1014 TTTCTTGTGGGGGCTGTATACAGCCAACTATATGTTAGTTACAACTGTGGAAAC 1073
Qy 1021 AATGATGTGAATATCTCCGCTCCAGAACCTTCCGCTCCAGAAATCTCCAGAGGC 1080
Db 1074 AATGATGTGAATATCTCCGCTCCAGAACCTTCCGCTCCAGAAATCTCCAGAGGC 1133
Qy 1081 AACGTTGTGACATATGATGTCACAGCAATGCAATCTCTATGAGCCTGTGTAATCCC 1140
Db 1134 AACGTTGTGACATATGATGTCACAGCAATGCAATCTCTATGAGCCTGTGTAATCCA 1193
Qy 1141 GACCATGTTGTTATTAATATGATGTGCTTACGTAAGGAGATAGCAAGAGCCATGAT 1200
Db 1194 GACCATGTTGTTATTAATATGATGTGCTTACGTAAGGAGATAGCAAGAGCCATGAT 1253
Qy 1201 GAGTACACTTCAGAAATATCAATGGGTGAAAGAACACATGTTTTCACAAACATGT 1260
Db 1254 GAGTACACTTCAGAAATATCAATGGGTGAAAGAACACATGTTTTCACAAACATGT 1313
Qy 1261 GAGATTCCTCTTTAGCTGCTCTTATATCTTGAATTTGATCTTGTGAGCTGAGC 1320
Db 1314 GAGATTCCTCTTTAGCTGCTCTTATATCTTGAATTTGATCTTGTGAGCTGAGC 1373

```
QY 1321 ACTAGATCCAGTTTAAAGCTGAAAATGAGGAAAAATCCACTCATTTCCACCAGTTGCT 1380
| | | | |
DB 1374 ACTAGATCGAGTTTAAAGCTGAAAATGAGGAAAAATCCACTCATTTCCACCAGTTGCT 1433
| | | | |
QY 1381 ACCATTCTGAGTATCTGACCAAGGCTCTCTGTTCCACCGGTACACCAATGGTGAAT 1440
| | | | |
DB 1434 ACCATCTGAGTATCTGACCAAGGCTCTCTGTTCCACCGGTACACCAATGGTGAAT 1493
| | | | |
QY 1441 GCATTCTCAAGAGCGGTCAATGCTGAAAACATATAGAGGCTTGTGGATTGGCC 1500
| | | | |
DB 1494 GCATTCTCAAGAGCGGTCAATGCTGAAAACATATAGAGGCTTGTGGATTGGCC 1553
| | | | |
QY 1501 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1533
| | | | |
DB 1554 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1586
| | | | |

RESULT 10
US-10-718-952-1
; Sequence 1, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-1

Query Match 95.5%; Score 1464.2; DB 19; Length 1760;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 AGTTTATGAGAAATTTTAAAGTAGAGTCCCTAATGTGAATGACCGAGACTGAGAT 60
| | | | |
DB 54 AGTTTCATGAGAAATTTTAAAGTAGAGTCCCTAATGTGAATGACCGAGACTGAGAT 113
| | | | |
QY 61 CAGTCCGTTGACAACTACGAAACACCGAACTTGTACGAGAAACAGAAATGGACCTAT 120
| | | | |
DB 114 CAGTCCGTTGACAACTACGAAACACCGAACTTGTGTTCAGAGAAACAGAAATGGACCTAT 173
| | | | |
QY 121 CAGTGAATGTCAAAACCAATCCGTCAATCAATTTTAAACCAACACCAATGTTCCA 180
| | | | |
DB 174 CAGTGAATGTCAAAACCAATCCGTCAATCAATTTTAAACCAACCAATGTTTCT 233
| | | | |
QY 181 AAATTGGGGTGATGCTTGTGGGTGGGGTGAAAACAAGGCTTACCTCACCGGTGAT 240
| | | | |
DB 234 AAATTAGGGGTATGCTTGTGGGTGGGGTGAAAACAAGGCTTACCTCACCGGTGAT 293
| | | | |
QY 241 GTTATGCTAACAGAGAGCATTTGATGGGCTTACAAAGACAAAGATTCAACAGCAAT 300
| | | | |
DB 294 GTTATGCTAACAGAGAGCATTTGATGGGCTTACAAAGACAAAGATTCAACAGCAAT 353
| | | | |
QY 301 TACTTTGCTCCCTCAAGGCTCAGCTATTGAGTTGATCTTTCAGAGGAGAGGAA 360
| | | | |
DB 354 TACTTTGCTCCCTCAAGGCTCAGCTATTGAGTTGATCTTTCAGAGGAGAGGAA 413
| | | | |
QY 361 ATCTATGCCCATTCAGAGTCTGCTTCCATGTTAATCTGACGACATTTGTTGGG 420
| | | | |
```

```
DB 414 ATCTATGCCCATTCAGAGTCTGCTTCCATGTTAATGTTAAACCTGACGACATTTGTGGG 473
| | | | |
QY 421 GGATGGATATCAGCAATGAACTGCTGATGCGCATGAGCGCAAGGCAAGGTGTTGAC 480
| | | | |
DB 474 GGATGGATATCAGCAATGAACTGCTGATGCGCATGAGCGCAAGGCAAGGTGTTGAC 533
| | | | |
QY 481 ATGATTTGCAAGACAGTTGAGGCTTTAATGAAATCATGTTTCCACTCCCGGAATC 540
| | | | |
DB 534 ATGATTTGCAAGACAGTTGAGGCTTTAATGAAATCATGTTTCCACTCCCGGAATC 593
| | | | |
QY 541 TACGACCCGGGATTTCAATGCTGCCAACAGAGGAGCGTGCAACAGTGTATTAAGGCG 600
| | | | |
DB 594 TATGACCCGGGATTTCAATGCTGCCAACAGAGGAGCGTGCAACAGTGTATTAAGGCG 653
| | | | |
QY 601 ACAAGCAAGAGCAAGTTTACGCAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 660
| | | | |
DB 654 ACAAGCAAGAGCAAGTTTACGCAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 713
| | | | |
QY 661 AAAGTGACAAAGTGTGTTCTGTTGACCTGCAACACAGAGGTATAGCAATTTGGTT 720
| | | | |
DB 714 AAAGTGACAAAGTGTGTTCTGTTGACCTGCAACACAGAGGTATAGCAATTTGGTT 773
| | | | |
QY 721 GTAGGCTTAAATGACACCATGAGAAATCTTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
| | | | |
DB 774 GTAGGCTTAAATGACACCATGAGAAATCTTTGGCTGCTGAGACAGAAATGAGGCTGAG 833
| | | | |
QY 781 ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGATGAGAAATGTTCTTTCTTAAT 840
| | | | |
DB 834 ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGATGAGAAATGTTCTTTCTTAAT 893
| | | | |
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTATCTTGGCATGCGAGAAACT 900
| | | | |
DB 894 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTATCTTGGCATGCGAGAAACT 953
| | | | |
QY 901 TTGATTTGAGAGATGACTTCAAGAGTGTGACAGACCAATGAATCTGTGTGGTAT 960
| | | | |
DB 954 TTGATTTGAGAGATGACTTCAAGAGTGTGACAGACCAATGAATCTGTGTGGTAT 1013
| | | | |
QY 961 TTTCTTGTGGGGCTGTATCAAGCCCAATCTATATAGTTAGTTACAAACATCTGGGAAC 1020
| | | | |
DB 1014 TTTCTTGTGGGGCTGTATCAAGCCCAATCTATATAGTTAGTTAGTTACAAACATCTGGGAAC 1073
| | | | |
QY 1021 AATGATGTAATGATCTCTGGCTTCAAAACCTTCGCTCCAGAAATCTCCAAAGAC 1080
| | | | |
DB 1074 AATGATGTAATGATCTCTGGCTTCAAAACCTTCGCTCCAGAAATCTCCAAAGAC 1133
| | | | |
QY 1081 AAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
| | | | |
DB 1134 AAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
| | | | |
QY 1141 GACCATGTTGTTGTTAATTAATGATGCTTACGTAAGGAGATAGCAAGAGCCATGAT 1200
| | | | |
DB 1194 GACCATGTTGTTGTTAATTAATGATGCTTACGTAAGGAGATAGCAAGAGCCATGAT 1253
| | | | |
QY 1201 GAGTACACTTCAAGATATTCATGCGGTGAAAAGAACACATTTGTTTGAACAACAATG 1260
| | | | |
DB 1254 GAGTACACTTCAAGATATTCATGCGGTGAAAAGAACACATTTGTTTGAACAACAATG 1313
| | | | |
QY 1261 GAGATTCCTCTTTTACGCTCTTATTAATCTTGAATGATGCTTCTTGTGAGGCTGAG 1320
| | | | |
DB 1314 GAGATTCCTCTCTTTTACGCTCTTATTAATCTTGAATGATGCTTCTTGTGAGGCTGAG 1373
| | | | |
QY 1321 ACTAGATCCAGTTTAAAGCTGAAAATGAGGAAAAATCCACTCATTTCCACCAGTTGCT 1380
| | | | |
DB 1374 ACTAGATCCAGTTTAAAGCTGAAAATGAGGAAAAATCCACTCATTTCCACCAGTTGCT 1433
| | | | |
QY 1381 ACCATTCTGAGTATCTGACCAAGGCTCTCTGTTCCACCGGTACACCAATGGTGAAT 1440
| | | | |
DB 1434 ACCATCTGAGTATCTGACCAAGGCTCTCTGTTCCACCGGTACACCAATGGTGAAT 1493
| | | | |
QY 1441 GCATTGTCAAGAGCGGTCAATGCTGAAAACATATAGAGGCTTGTGGATTGGCC 1500
| | | | |
DB 1494 GCATTGTCAAGAGCGGTCAATGCTGAAAACATATAGAGGCTTGTGGATTGGCC 1553
| | | | |
```

QY 1501 CCAGAGATTAACATGATTCCTCGAGTACAGTGA 1533
DB 1554 CCAGAGATTAACATGATTCCTCGAGTACAGTGA 1586

RESULT 11

US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-5

Query Match 95.4%; Score 1462.6; DB 14; Length 1533;

Best Local Similarity 97.1%; Pred. No. 0;

Matches 1489; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1 AGTTCATGAGAAATTTAAAGTAGAGAGTCCCTAATGTAAGTACACCGAGCTGAGATT 60
DB 1 AGTTCATGAGAAATTTAAAGTAGAGTGTCTTAATGTAAGTACACCGAGCTGAGATT 60
QY 61 CAGTCCGCTGACAACTACGAAACCAACGAACTTTGTTACGAGAAACAGAAATGGACCTAT 120
DB 61 CAGTCCGCTGACAACTACGAAACCAACGAACTTTGTTACGAGAAACAGAAATGGACCTAT 120
QY 121 CAGTGGATTGTCAAAACCAATCTGCTCAATCAAAATTTAAACCAACCAACCTATTTCA 180
DB 121 CAGTGGATTGTCAAAACCAATCTGCTCAATCAAAATTTAAACCAACCAATCATGTTCTT 180
QY 181 AAATTGGGGGTGATGTTTGGGTGGGTGGGTAACAAACGGCTCTACCCCTCAACGGTGGT 240
DB 181 AAATTGGGGGTGATGTTTGGGTGGGTGGGTAACAAACGGCTCTACCCCTCAACGGTGGT 240
QY 241 GTTATTGCTAACAGAGAGACATTTTCATGGGCTAACAAAGAACAAAGATTCAACAGCAAT 300
DB 241 GTTATTGCTAACAGAGAGAGACATTTTCATGGGCTAACAAAGAACAAAGATTCAACAGCAAT 300
QY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCGAGTTGGGTCCTTCCAGGGAAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCGAGTTGGGTCCTTCCAGGGAAGAGAA 360
QY 361 ATCTATGCCCCCATTAAGAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCCCATTAAGAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGG 420
QY 421 GGATGGATATACGAAACATGAACCTGAGTGAATGCGATGGGCAAGGAGGATTGGAC 480
DB 421 GGATGGATATACGAAACATGAACCTGAGTGAATGCGATGGGCAAGGAGGATTGGAC 480
QY 481 ATCGATTTGACAGAGAGGTTGAGGCTTTAATGGAATCCATGATTCCTACCTCCCGAATC 540
DB 481 ATCGATTTGACAGAGAGGTTGAGGCTTTAATGGAATCCATGATTCCTACCTCCCGAATC 540
QY 541 TACGACCCCGGATTTCACTTGTCTGCCAAACAAAGAGAGCGTGCACAAACGTGATTAAGGCG 600

DB 541 TACGACCCCGGATTTCACTTGTCTGCCAAACAAAGAGAGCGTGCACAAACGTGATTAAGGCG 600
QY 601 ACAAAGCAAGAGCAAGTTGAGCAAAATCATCAAAAGCATCAAGGCTTTAAGAAAGCACCC 660
DB 601 ACAAAGCAAGAGCAAGTTGAGCAAAATCATCAAAAGCATCAAGGCTTTAAGAAAGCACCC 660
QY 661 AAAGTGCACAAAGTGTGCTGCTGAGACTGGCAACAGAGAGGATATGCAATTTGGTT 720
DB 661 AAAGTGCACAAAGTGTGCTGCTGAGACTGGCAACAGAGAGGATATGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTCTGGCTGCTGTGAGACAAATAGGCTGAG 780
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTCTGGCTGCTGTGAGACAAATAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGTATGCAATTCCTGCTGTGATGAAATGTTCTTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTGTATGCAATTCCTGCTGTGATGAAATGTTCTTTCAATTAAT 840
QY 841 GGAAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACAT 900
DB 841 GGAAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACAT 900
QY 901 TTGATTTGGTGAATGACTTCAAGAGTGGTACAGCAAAATGAATCTGTGTTGGTGAAT 960
DB 901 TTGATTTGGTGAATGACTTCAAGAGTGGTACAGCAAAATGAATCTGTGTTGGTGAAT 960
QY 961 TTTCTTGTGGGGGCTGTATACAGCCAAATCTAATAGTTAGTTACAAACCATCTGGGAAAC 1020
DB 961 TTTCTTGTGGGGGCTGTATACAGCCAAATCTAATAGTTAGTTACAAACCATCTGGGAAAC 1020
QY 1021 AATGATGATGATATCTCTGGCTCCAAACCTTCCGCTCCAGAGAAATCTTCCAAAGAGC 1080
DB 1021 AATGATGATGATATCTCTGGCTCCAAACCTTCCGCTCCAGAGAAATCTTCCAAAGAGC 1080
QY 1081 AAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 AAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GACCATGTTGTTGTTAATTAAGTATGCTTACGTAGGGGATAGCAAGAGACCATGAT 1200
DB 1141 GACCATGTTGTTGTTAATTAAGTATGCTTACGTAGGGGATAGCAAGAGACCATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGCGGTGAAAGAACACCATTTGTTGCAACACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGCGGTGAAAGAACACCATTTGTTGCAACACATGT 1260
QY 1261 GAGATTTCCCTTTAGCTGCTCTAATTAATCTTGAATGCTGCTGCTGAGCTGAGC 1320
DB 1261 GAGATTTCCCTTTAGCTGCTCTAATTAATCTTGAATGCTGCTGCTGAGCTGAGC 1320
QY 1321 ACTAAGATTCAGTTTAAAGCTGAAATAGGGAATTCATCTCAATCCACCACTGCT 1380
DB 1321 ACTAAGATTCAGTTTAAAGCTGAAATAGGGAATTCATCTCAATCCACCACTGCT 1380
QY 1381 ACCATTCCTAGCTTACCTAAGCAAGGCTCTCTGTTTCAACCGGATCAACAGTGTGAAT 1440
DB 1381 ACCATTCCTAGCTTACCTAAGCAAGGCTCTCTGTTTCAACCGGATCAACAGTGTGAAT 1440
QY 1441 GATTTGTCAAAAGCAAGCGTGAATGCTGGAATAAATGAAGGCTTTGTTGATTTGGCC 1500
DB 1441 GATTTGTCAAAAGCAAGCGTGAATGCTGGAATAAATGAAGGCTTTGTTGATTTGGCC 1500
QY 1501 CCAGAGAAATTAACATGATTCCTGAGTACAAAGTGA 1533
DB 1501 CCAGAGAAATTAACATGATTCCTGAGTACAAAGTGA 1533

RESULT 12

US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:

APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streif, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 5
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-718-952-5

Query Match 95.4%; Score 1462.6; DB 19; Length 1533;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1489; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

1 ATGTCATCGAAGATTTTAAAGTAGAGAGTCCATATGTAGATACACCGAGACTGAGATT 60
1 AGTTCATGAGATTTTAAAGTTGAGTGTCTTAATGTAATGACCGAGACTGAGATT 60
61 CAGTCCGTTTACACTACGAAACACCGAACTTGTTCACGAGACAGGAATGACACTTAT 120
61 CAGTCCGTTTACACTACGAAACACCGAACTTGTTCACGAGACAGGAATGACACTTAT 120
121 CAGTGGATTGTAAACCCAAATCCGTCACTACATTTAAACCAACCCATGTTTCA 180
121 CAGTGGATTGTAAACCCAAATCCGTCACTACATTTAAACCAACCCATGTTTCT 180
181 AAATTGGGGGTGATGCTTGGGTGGGGTGGAAACACGCGCTTACCCCTCAACCGGTGT 240
181 AAATTGGGGGTGATGCTTGGGTGGGGTGGAAACACGCGCTTACCCCTCAACCGGTGT 240
241 GTTATTGCTTAAACAGAGAGACATTTATGAGGCTTAAACAGACAGATTTAAACAGCAAT 300
241 GTTATTGCTTAAACAGAGAGACATTTATGAGGCTTAAACAGACAGATTTAAACAGCAAT 300
301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGAGTCTTCCAGGGAGAGAA 360
301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGAGTCTTCCAGGGAGAGAA 360
361 ATCTATGCCCATTTCAAGAGTGTCTTCCATGTTTATCTGACGACATTTGTTTGGG 420
361 ATCTATGCCCATTTCAAGAGTGTCTTCCATGTTTATCTGACGACATTTGTTTGGG 420
421 GAGTGGGATATACGACACATGAACCTGGCTGATGCTGATGCGGCAAGGCAAGTGTTCAC 480
421 GAGTGGGATATACGACACATGAACCTGGCTGATGCTGATGCGGCAAGGCAAGTGTTCAC 480
481 ATCGATTTCAGAGAGATTTAGGCTTACATGGAATTCATGTTTCCATCCCGGGAATC 540
481 ATCGATTTCAGAGAGATTTAGGCTTACATGGAATTCATGTTTCCATCCCGGGAATC 540
541 TAGACCCCGGATTTTCTGCTGCAACCAAGAGACGCTGCCAACACGTTGATTAAGAGGC 600
541 TAGACCCCGGATTTTCTGCTGCAACCAAGAGACGCTGCCAACACGTTGATTAAGAGGC 600
601 ACAAGCAAGAGACAGTTCAGCAATCATCAAAAGACATCAAGGCGTTTAAAGAGCAACC 660
601 ACAAGCAAGAGACAGTTCAGCAATCATCAAAAGACATCAAGGCGTTTAAAGAGCAACC 660
661 AAAGTGGACAGAGTGTGTCTGCTGAGTGTGCAACACAGAGAGATTAAGCAATTTGGTT 720
661 AAAGTGGACAGAGTGTGTGTCTGCTGAGTGTGCAACACAGAGAGATTAAGCAATTTGGTT 720

721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
781 ATTTCCTTCCACTTGTATGCAATGCTGCTGTGATGAGAAATGTTCTTTCTTAAT 840
781 ATTTCCTTCCACTTGTATGCAATGCTGCTGTGATGAGAAATGTTCTTTCTTAAT 840
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGGCCATGCGAGAAACT 900
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGGCCATGCGAGAAACT 900
901 TTGATTTGATGATGATGATCTTCAAGAGTGTGATGATGATGATGATGATGATGAT 960
901 TTGATTTGATGATGATGATCTTCAAGAGTGTGATGATGATGATGATGATGATGAT 960
961 TTTCTTGGGGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
961 TTTCTTGGGGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1081 AACTGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1081 AACTGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1141 GACCAATGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1141 GACCAATGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1201 GAGTACACTTCAAGATATTTATGAGTGGTGAAGAAACACATTTGTTTTCACAAACATGT 1260
1201 GAGTACACTTCAAGATATTTATGAGTGGTGAAGAAACACATTTGTTTTCACAAACATGT 1260
1261 GAGGATTCCTTTTATGCTGCTCTATTTATCTTGAATGTTGCTTCTTCTGAGCTGAGC 1320
1261 GAGGATTCCTTTTATGCTGCTCTATTTATCTTGAATGTTGCTTCTTCTGAGCTGAGC 1320
1321 ACTAATTCAGTTTAAAGCTGAAGAAATGAGGAAATTCATCTCACTTCCACCCAGTTGCT 1380
1321 ACTAATTCAGTTTAAAGCTGAAGAAATGAGGAAATTCATCTCACTTCCACCCAGTTGCT 1380
1381 ACCATTCCTCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGCTTACCAAGTGTGAT 1440
1381 ACCATTCCTCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGCTTACCAAGTGTGAT 1440
1441 GCATTGTCAAAAGCAGCGTGAATGCTGGAATAATGAAGAGGCTTGTGATTTGGCC 1500
1441 GCATTGTCAAAAGCAGCGTGAATGCTGGAATAATGAAGAGGCTTGTGATTTGGCC 1500
1501 CCAGGAATTAACATGATTTCTGAGTACAGTGA 1533
1501 CCAGGAATTAACATGATTTCTGAGTACAGTGA 1533

RESULT 13
US-10-424-599-70167
Sequence 70167, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223) B
CURRENT APPLICATION NUMBER: US/10/424,599
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 70167

LENGTH: 1989
 TYPE: DNA
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1
 US-10-424-599-70167

Query Match 95.0%; Score 1456.4; DB 18; Length 1989;
 Best Local Similarity 97.3%; Pred. No. 0;
 Matches 1492; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

1 ATGTCATCGAGAAATTTAAAGTAGAGAGTCTAATAGTAGAGACCGAGACTGAGATT 60
 217 ATGTCATCGAGAAATTTAAAGTAGAGTGTCTAATAGTAGAGACCGAGACTGAGATT 276
 61 CAGTCGCTGATCACTACGAAACCAACGAACTTGTTCACGAGAACGAAATGGACCTAT 120
 277 CAGTCGCTGATCACTACGAAACCAACGAACTTGTTCACGAGAACGAAATGGACCTAT 336
 121 CAGTCGATGTCAAAACCCAAATCCGTCATCACTAATTTAAACCAACCCGATGTTCA 180
 337 CAGTCGATGTCAAAACCCAAATCCGTCATCACTAATTTAAACCAACCCGATGTTCT 396
 181 AAATGGGGGTGATGCTTGAGGT-GGGGTGAACAACGCTCTACCTCACCGGTG 239
 397 AAATGGGGGTATGCTTGAGGTGGGGTGAACAACGCTCTACCTCACCGGTG 456
 240 TGTATGTCTAACAGAGAGACATTTTCATGGGCTACAAAGACAGATTTCAACAGCAA 299
 457 TGTATGTCTAACAGAGAGAGACATTTTCATGGGCTACAAAGACAGATTTCAACAGCAA 516
 300 TTAATTTGGCTCCCTCAACCAAGCTCAGTATTCAGATTGATCCTTCAAGGAGAGA 359
 517 TTAATTTGGCTCCCTCAACCAAGCTCAGTATTCAGATTGATCCTTCAAGGAGAGA 576
 360 AATCTATGCCCATTTCAAGAGTGTCTTCAATGTTAATCCGACGACATTTGTTGG 419
 577 AATCTATGCCCATTTCAAGAGTGTCTTCAATGTTAATCCGACGACATTTGTTGG 636
 420 GGGATGGATATCAGAACATGAACCTGCTGATGTCATGGCCAGGCAAAAGTGTGA 479
 637 GGGATGGATATCAGAACATGAACCTGCTGATGTCATGGCCAGGCAAAAGTGTGA 696
 480 CATCGATTTGCAAGACAGTGAAGCTTTCATGATTCATGTTCCACTCCCGGAAT 539
 697 CATCGATTTGCAAGACAGTGAAGCTTTCATGATTCATGTTCCACTCCCGGAAT 756
 540 CTACGACCCGGAATTTCAATGCTGCAACCAAGAGAGCGTCCCAACAGGTGTTAAGG 599
 757 CTATGACCCGGAATTTCAATGCTGCAACCAAGAGAGCGTCCCAACAGGTGTTAAGG 816
 600 CACAAAGCAAGAGCAAGTTCAGCAATCAACAAGACATCAAGCGCTTTAAGAACCCAC 659
 817 CACAAAGCAAGAGCAAGTTCAGCAATCAACAAGACATCAAGCGCTTTAAGAACCCAC 876
 660 CAAAGTGAACAAGTGTGTCTCTGTGACTGCCACACAGAGAGGTATAGCAATTTGGT 719
 877 CAAAGTGAACAAGTGTGTCTCTGTGACTGCCACACAGAGAGGTATAGCAATTTGGT 936
 720 TGTAGGCTTTAATGACACCATGAGAAATCTTGGGCTGTGAGACAGAAATGAGGCTGA 779
 937 TGTGGGCTTTAATGACACCATGAGAAATCTTGGGCTGTGAGACAGAAATGAGGCTGA 996
 997 GATTTCTCTTCCACTTGTATGTCATGCTGTTGTTATGAAAAATGTTCTTTCACTTAA 1056
 840 TGAAGCCCTCAGAACACTTTTGTATCAGAGGCTGATTTGATCTTGCATGCCAGAACAC 899
 1057 TGAAGCCCTCAGAACACTTTTGTATCAGAGGCTGATTTGATCTTGCATGCCAGAACAC 1116
 900 TTGATGTGTGAGATGATCTTCAAGAGGTGTCAGCAAAATGAATCTGTGTGTTGA 959

1117 TTGATGTGTGAGATGATCTTCAAGAGGTGTCAGACCAAAATGAATCTGTGTGTTGA 1176
 960 TTTTCTTTGGGGGCTGGTATCAAGCCAACTATAGTTAGTTACACACTTGGGAAA 1019
 1177 TTTCTTTGGGGGCTGGTATCAAGCCAACTATAGTTAGTTACACACTTGGGAAA 1236
 1020 CAATGATGATGAATCTCGGGCTCACAACCTTCGGCTCCAGGAATCTCCAAAG 1079
 1237 CAATGATGATGAATCTTCGGCTCACAACCTTCGGCTCCAGGAATCTCCAAAG 1296
 1080 CAAGTTGTGAAGATATGATGATCAAGCAATGCTCTATGAGCTGTGAAATCC 1139
 1297 CAAGTTGTGAAGATATGATGATCAAGCAATGCTCTATGAGCTGTGAAATCC 1356
 1140 CAGCAATGTTGTTTATTAATGATGCTTACGTAAGGGGATAGCAAGAGCCATGGA 1199
 1357 AGACCAATGTTGTTTATTAATGATGCTTACGTAAGGGGATAGCAAGAGCCATGGA 1416
 1200 TGAATCACTGAGATATGATGATGAGGAGAAAGAACCAATGTTTGCACAACATG 1259
 1417 TGAATCACTGAGATATGATGATGAGGAGAAAGAACCAATGTTTGCACAACATG 1476
 1260 TGAATCACTGAGATATGATGATGAGGAGAAAGAACCAATGTTTGCACAACATG 1319
 1477 CAGGATTCCTCTTATAGTGTCTTATTAATCTTGAATGATGCTTCTTGTGAGCTCAG 1536
 1320 CACTGATATCCAGTTTAAAGCTGAAATAGAGGAAATTCCTCATTTCCACCAGTTGC 1379
 1537 CACTGATATCCAGTTTAAAGCTGAAATAGAGGAAATTCCTCATTTCCACCAGTTGC 1596
 1380 TACCAATTCACAGTATTCGACCAAGGCTCTGGTTCACCGGGTACACAGTGTGAA 1439
 1597 TACCAATTCACAGTATTCGACCAAGGCTCTGGTTCACCGGGTACACAGTGTGAA 1656
 1440 TGCATTTGCAAGACAGCGTGCATGCTGGAACCAATATGAGGCTGTGTGATGAGC 1499
 1657 TGCATTTGCAAGACAGCGTGCATGCTGGAACCAATATGAGGCTGTGTGATGAGC 1716
 1500 CCCAGGAATTAACATGATTTCTGAGTACAGTGA 1533
 1717 CCCAGGAATTAACATGATTTCTGAGTACAGTGA 1750

RESULT 14
 US-10-424-599-12022
 Sequence 12022, Application US/10424599
 Publication No. US20040031072A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa Thomas J
 APPLICANT: Kovalic David K
 APPLICANT: Zhou Yihua
 APPLICANT: Cao Yongwei
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53223) B
 CURRENT APPLICATION NUMBER: US/10/424, 599
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 285684
 SEQ ID NO 12022
 LENGTH: 2018
 TYPE: DNA
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1
 US-10-424-599-12022

Query Match 82.3%; Score 1261; DB 18; Length 2018;
 Best Local Similarity 88.9%; Pred. No. 0;
 Matches 1363; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

1 ATGTCATCGAGAAATTTAAAGTAGAGAGTCTAATAGTAGAGACCGAGACTGAGATT 60
 98 ATGTCATCGAGAGTTTCAAGGTGAGAGTCTTAACGTGAAGTACACAGAGACTGAGATT 157

```
OY 61 CAGTCCGTGTAACAATGACCAACCAAGCTTGTTCAGAGAGACAGATGGACCTAT 120
DB 158 CAGTCCGTGTAACAATGACCAACCAAGCTTGTTCAGAGAGACAGATGGACCTAT 217
OY 121 CAGTGGATTGTCAAAACCAAAATCCGTCAATCAATTTTAAACCAACCAATGTTTCA 180
DB 218 CAGTGGATTGTCAAAACCAAAATCCGTCAATCAATTTTAAACCAACCAATGTTTCA 277
OY 181 AAATTGGGGGTGATGTTTGGGTTGGGGTGGAAAACAAGGCTTACCTTCAACCGTGGT 240
DB 278 AAACCTAGGGGTATGTTGTGGGGTGGGGGAAAACAATGGCTCAACCTTCACTGGTGT 337
OY 241 GTTATTGTCTAAGAGAGACATTTTCATGGGCTCAAAAGACAAAGATTCAAGACCAAT 300
DB 338 GTTATTGTCTAAGAGAGGATTTTCAATGGGCAACAAAGACAAAGATTCAAGACCAAT 397
OY 301 TACTTTGGCTCCCTCAACCCCAAGCTCAAGTATTTCAGATTGGATCTTTCAGAGGAGAGAA 360
DB 398 TACTTTGGGTCCCTGACCCAGAGCATCAGCTATCAGATGGGGTCTTTCCAAAGGGAGAA 457
OY 361 ATCTAATCCCAATGAGAGTCTGCTTCAATGGTTAACTCTGACGACATTTGTTTGGG 420
DB 458 AATATATGCTCCATTCAGAGCCCTGCTTCCAAATGGTAACTCTGATGATGTTGTTTGG 517
OY 421 GGATGGGATATCAGCAATGAACTGGCTGATGCGATGGCCAGAGGCAAGGTTTGGAC 480
DB 518 GGATGGGATATCAGTAACTTGAACCTGGCTGATGCGATGGCCAGAGGCAAGGTTTGGAT 577
OY 481 ATCGATTTCAGAGAGATTGAGGCTTTCATGGAATTCATGAGTTTCACTCCCGGAATC 540
DB 578 ATCGACCTCAGAGAAACAGTTGAGGCGGTACATGGAATTCATGAGTTTCACTCCCGGAATC 637
OY 541 TACGACCCGAGTTTCTTCTGCTGCAACCAAGAGAGAGCTGCAACCAAGTGAATTAAGGCG 600
DB 638 TATGACCCGAGTTTCTTCTGCTGCAACCAAGAGAGAGCTGCAACCAAGTGAATTAAGGCG 697
OY 601 ACAAGCAAGAGCAAGTTCAGCAAAATCATCAAAAGACATCAAGGCTTTTAAAGAGGACAC 660
DB 698 ACAAGAGAGAGCAAGTTCAGCAAAATCATCAAAAGACATTAAGGAGTTTCAAGAGAGACAT 757
OY 661 AAAGTGACAGAGGTGTTTCTGCTGCAACCAAGAGAGAGTGAATTAAGCAATTTGGTT 720
DB 758 AAAGTTGACAGAGGTGTTTCTGCTGCAACCAAGAGAGAGTGAATTAAGCAATTTGGTT 817
OY 721 GTAGGCTTAATGACCAATGAGGATCTTTGGCTGCTGGAACAAGATTAAGGCTGAG 780
DB 818 GTGGGACTAAACGACCAATGAGAAACCTTTGCTTCTTGGACAGGAATGAGGCTGAG 877
OY 781 ATTTCTCTTCCACCTTGTATGCGATTTGCTGTGATGGAATAATTTCTTTTCAATTAAT 840
DB 878 ATTTCTCTTCCACCTTGTATGCGATTTGCTGTGATGGAATAATTTCTTTTCAATTAAT 937
OY 841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAGAACT 900
DB 938 GGAAGCCCTCAGAAACCTTTTGTCTCAGAGATCTGATGATCTGGCATCAAAAGGAATAGT 997
OY 901 TTGATTGGTGGAGATGATCTTCAAGAGTGGTCAAGCAAAATGAAATCTGTGTTGGTGAAT 960
DB 998 TTGATTGGGAGAGATGATCTTAAAGTGTGCAAGCAAAATGAAATCGGTGTTGGTGAAT 1057
OY 961 TTTCTTGGGGGCTGTGTCAAGCAACATCTATAGTTAGTTTAAACCAACCTCTGGGAAAC 1020
DB 1058 TTTCTTGGGGGCTGTGTCAAGCAACATCTATAGTTAGTTTAAACCAACCTCTGGGAAAC 1117
OY 1021 AATGATGATGATATCTCTGCGCTCAACAAACCTTCCGCTCAAGGAAATCTTCAAGAC 1080
DB 1118 AATGATGATGATATCTCTCAAGCCCTCAAACTTCCGCTCAAGGAAATCTTCAAGAC 1177
OY 1081 AAGCTTGGAGATATGTTCAACAGCAATGCGATCTCTATAGAGCTGTGAAACATCCC 1140
DB 1178 AAGTGTGTGAGATATGTTCAACAGCAAGCGCATCTCTATAGAGCTGTGAAACATCCC 1237
```

```
OY 1141 GACCATGTTGTTGTTAATTAAGTATGTGCTTACGTAGGGGATAGCAGAGACCATGAT 1200
DB 1238 GACCATGTTGTTGTTAATTAAGTATGTACGTAGGGGACAGAGAGGCGCATGAT 1297
OY 1201 GAGTACACTTCAAGAGATATTCATGGGTGGAAGAACACATTTGTTTTCACAAACATGT 1260
DB 1298 GAGTACCTTTCTGAGATATTTCAATGGGCGGAAAGAACAAATGTTGTTTCAACAACTTGT 1357
OY 1261 GAGATTCCTTTTATGCTGCTCTTATTAATCTTGAACCTTGTCTTCTGAGAGCTGAGC 1320
DB 1358 GAGATTCCTTATGCTGCTCTTATTAATCTTGAACCTTGTCTTCTGAGAGCTTATG 1417
OY 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTTCACTCATTCACCCAGTTGCT 1380
DB 1418 ACTGAAATCGAGTTTAAAGCTGAAATGAGGAAATTTCACTCATTCACCCAGTTGCT 1477
OY 1381 ACCATTTCAGGTATCTGACCAAGGCTCTCTGTTTCCACCGGGATACACAGTGGTAAAT 1440
DB 1478 ACCATTCCTCAGTTTAACTTCAACCAAGGCTCTCTGTTTCCACCGAGTACGCGAGTGGTAAAT 1537
OY 1441 GCATTGTCAAAGCAGCTGCAATGCTGGAATAACATTAATGAGGCTTGTGTTGATTTGCC 1500
DB 1538 GCATTGTCAAAGCAGCTGCAATGCTGGAATAACATTAATGAGGCTTGTGTTGATTTACT 1597
OY 1501 CCAGAGAAATTAATGATTTCTGAGATTAAGTGA 1533
DB 1598 CCTGAGAAACAAATGATCTCTGAGATTAAGTGA 1630
```

```
RESULT 15
US-10-424-599-12021/c
; Sequence 12021, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12021
; LENGTH: 2582
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1
US-10-424-599-12021
```

```
Query Match 76.7%; Score 1176.4; DB 18; Length 2582;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1183; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 AATTCATCGAGAAATTTTAAAGTATGAGAGTCTTAATGTGAATGACGAGCTGAGATT 60
DB 2454 AATTCATCGAGAAATTTTAAAGTATGAGAGTCTTAATGTGAATGACGAGCTGAGATT 2395
OY 61 CAGTCCGTGTAACAATGACCAACCAAGCTTGTTCAGAGAGACAGATGGACCTAT 120
DB 2394 CAGTCCGTGTAACAATGACCAACCAAGCTTGTTCAGAGAGACAGATGGACCTAT 2335
OY 121 CAGTGGATTGTCAAAACCAAAATCCGTCAATCAATTTTAAACCAACCAATGTTTCA 180
DB 2334 CAGTGGATTGTCAAAACCAAAATCCGTCAATCAATTTTAAACCAACCAATGTTTCA 2275
OY 181 AAATTGGGGGTGATGTTTGGGTTGGGGTGGAAAACAAGGCTTACCTTCAACCGTGGT 240
DB 2274 AAATTGGGGGTGATGTTTGGGTTGGGGTGGAAAACAAGGCTTACCTTCAACCGTGGT 2215
OY 241 GTTATTGTCTAAGAGAGACATTTTCATGGGCTACAAAGACAAAGATTCAACCAAGCAAT 300
```

```
Db 2214 GTTATTGCTAAGAGAGGCAATTTTCATGGGCTACAAAGSACAAAGATTCACAAAGCCAAAT 2155
Qy 301 TACTTTGGGCTCCCTCAACCCAAAGCCTCAAGCTATTGAGTTGGATCTTCCAGGGAGAGAA 360
Db 2154 TACTTTGGGCTCCCTCAACCCAAAGCCTCAAGCTATTGAGTTGGATCTTCCAGGGAGAGAA 2095
Qy 361 ATCTATGCCCCCATTCAAAGAGTGTGCTTCCAAATGGTTAAATCCATGAGCATATGTTGGTGG 420
Db 2094 ATCTATGCCCCCATTCAAAGAGTGTGCTTCCAAATGGTTAAATCCATGAGCATATGTTGGTGG 2035
Qy 421 GGAATGGGAATATCAGCAACATGAACCTGCTGATGCCATGGCCAGGGCAAAAGTGTGGAC 480
Db 2034 GGAATGGGAATATCAGCAACATGAACCTGCTGATGCCATGGCCAGGGCAAAAGTGTGGAC 1975
Qy 481 ATCGATTTCGAGAGAGAGATTGAGGCTTACATGAAATCCATGGTTCCACTCCCGGAATC 540
Db 1974 ATCGATTTCGAGAGAGAGATTGAGGCTTACATGAAATCCATGGTTCCACTCCCGGAATC 1915
Qy 541 TACGACCCGGAATTTCTATGCTGCAACCAAGAGAGCGTCCCAACACGTGATTAAGGCG 600
Db 1914 TACGACCCGGAATTTCTATGCTGCAACCAAGAGAGCGTCCCAACACGTGATTAAGGCG 1855
Qy 601 ACAAGACAAGAGCAAGTTCAGCAATCATCAAAAGACATCAAGSGCTTTTAAAGAAAGCCACC 660
Db 1854 ACAAGACAAGAGCAAGTTCAGCAATCATCAAAAGACATCAAGSGCTTTTAAAGAAAGCCACC 1795
Qy 661 AAAGTGACAAGGTGTTGCTCTGTGGAATGCAACAGAGAGATTAAGCAATTTGGTT 720
Db 1794 AAAGTGACAAGGTGTTGCTCTGTGGAATGCAACAGAGAGATTAAGCAATTTGGTT 1735
Qy 721 GTAAGGCTTAATGACCAATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
Db 1734 GTAAGGCTTAATGACCAATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 1675
Qy 781 ATTTCCCTTCCACCTTGTATGCAATGCTGCTGTGATGGAAAAATGTTCCCTTCATTAAAT 840
Db 1674 ATTTCCCTTCCACCTTGTATGCAATGCTGCTGTGATGGAAAAATGTTCCCTTCATTAAAT 1615
Qy 841 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATTTGATCTTGCCATCGGAGAAACAT 900
Db 1614 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATTTGATCTTGCCATCGGAGAAACAT 1555
Qy 901 TTGATTTGGAGATGATCTTCAAGAGTGGTCAAGCCAAATGAAATCTGTGTTGGTTGAT 960
Db 1554 TTGATTTGGAGATGATCTTCAAGAGTGGTCAAGCCAAATGAAATCTGTGTTGGTTGAT 1495
Qy 961 TTTCTTGTGGGGCTGGTATCAAGCCAAATCTATAGTTACAAACCATCTGGGAAAC 1020
Db 1494 TTTCTTGTGGGGCTGGTATCAAGCCAAATCTATAGTTACAAACCATCTGGGAAAC 1435
Qy 1021 AATGATGTATGAATCTCTCGGCTCCAAACCTTCGCTCCAAAGAAAATCTTCAAGAGC 1080
Db 1434 AATGATGTATGAATCTCTCGGCTCCAAACCTTCGCTCCAAAGAAAATCTTCAAGAGC 1375
Qy 1081 AATGATGTATGAATCTCTCGGCTCCAAACCTTCGCTCCAAAGAAAATCTTCAAGAGC 1140
Db 1374 AATGATGTATGAATCTCTCGGCTCCAAACCTTCGCTCCAAAGAAAATCTTCAAGAGC 1315
Qy 1141 GACCATGTTGTTATTAAGTATGTCCTTAAGTGGGGAATGCAAGAGAGCC 1194
Db 1314 GACCATGTTGTTATTAAGTATGTCCTTAAGTGGGGAATGCAAGAGAGCC 1261
```

Search completed: June 8, 2005, 11:26:00
Job time : 958.946 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 21:02:47 ; Search time 262.095 Seconds
(without alignments)
9570.620 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 agctcaccgagacattttaa.....tgatctcgcagacaaagta 1533

Scoring table: IDENTITY_NUC

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984.2	64.2	1959	4	US-09-727-628-1
2	979.4	63.9	1931	3	US-09-118-442-10
3	979.4	63.9	1931	3	US-09-677-064-10
4	384.8	25.1	1578	4	US-09-248-796A-3131
5	349.4	22.8	1231	3	US-09-397-787-34
6	349.4	22.8	1605	4	US-09-734-237B-74
7	334.6	21.8	1602	4	US-09-734-237B-72
8	161.4	10.5	3546	3	US-09-118-442-15
9	161.4	10.5	3546	3	US-09-677-064-15
10	159	10.4	3546	3	US-09-118-442-14
11	159	10.4	3546	3	US-09-677-064-14
12	92.6	6.0	77626	4	US-09-949-016-12608
13	63.8	4.2	294	4	US-09-313-294A-4684
14	44.2	2.9	1335	4	US-09-902-540-5256
15	44.2	2.9	34316	4	US-09-902-540-1257
16	38.8	2.5	7218	1	US-08-232-463-14
17	37	2.4	265038	4	US-09-949-016-15779
18	36.8	2.4	1141	4	US-09-806-708B-22
19	36.2	2.4	505	4	US-09-621-976-15639
20	36.2	2.4	568	4	US-09-370-767-2813
21	36.2	2.4	568	4	US-09-370-767-18095
22	35.6	2.3	1446	4	US-09-502-540-5772
23	35.6	2.3	72704	4	US-09-902-540-1273
24	35.4	2.3	274	3	US-09-118-442-21
25	35.4	2.3	274	3	US-09-677-064-21
26	35.4	2.3	3049	4	US-09-300-958A-42
27	35.4	2.3	3049	4	US-09-949-016-991

28	35.4	2.3	3068	4	US-09-949-016-1297	Sequence 1297, Ap
29	35.2	2.3	57299	4	US-09-949-016-12141	Sequence 12141, A
30	35.2	2.3	153866	4	US-09-949-016-16919	Sequence 16919, A
31	35	2.3	315	4	US-09-270-767-1779	Sequence 1779, Ap
32	35	2.3	315	4	US-09-270-767-17061	Sequence 17061, A
33	35	2.3	600	4	US-09-669-751-171	Sequence 171, Ap
34	35	2.3	2025	4	US-09-107-532A-2547	Sequence 2547, Ap
35	34.8	2.3	1915	4	US-09-902-540-7256	Sequence 7256, Ap
36	34.8	2.3	6744	4	US-09-902-540-682	Sequence 682, App
37	34.6	2.3	1098	3	US-09-221-017B-928	Sequence 928, App
38	34.4	2.2	4085	3	US-09-165-240-5	Sequence 5, Appl1
39	34.4	2.2	4085	3	US-08-568-059-5	Sequence 5, Appl1
40	34.4	2.2	1664976	4	US-08-916-421B-1	Sequence 1, Appl1
41	34.4	2.2	1664976	4	US-09-692-570-1	Sequence 1, Appl1
42	34	2.2	1422	4	US-09-450-209-3	Sequence 3, Appl1
43	34	2.2	2941	4	US-09-949-016-2167	Sequence 2167, Ap
44	34	2.2	112112	4	US-09-949-016-15639	Sequence 15639, A
45	33.8	2.2	1362	4	US-09-328-352-116	Sequence 116, App

ALIGNMENTS

```
RESULT 1
US-09-727-628-1
; Sequence 1, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727, 628
; PRIOR APPLICATION NUMBER: 2000-12-01
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)..(1699)
US-09-727-628-1

Query Match      64.2%; Score 984.2; DB 4; Length 1959;
Best Local Similarity 77.6%; Pred. No. 0;
Matches 1190; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

QY 1 AGTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTCAAGTACACGAGCTGAGATT 60
DB 137 AGTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTCAAGTACACGAGCTGAGATT 196
QY 61 CAGTCGAGTGTACACCTGACCAACCAACCACTGTTTACAGAACGAGAAATGACACTAT 120
DB 197 GAGTCGAGTGTACACCTGACCAACCAACCACTGTTTACAGAACGAGAAATGACACTAT 256
QY 121 CAGTCGAGTGTACACCTGACCAACCAACCACTGTTTACAGAACGAGAAATGACACTAT 180
DB 257 CAGTCGAGTGTACACCTGACCAACCAACCACTGTTTACAGAACGAGAAATGACACTAT 316
QY 181 AAATTGGGGGTATGCTTGTGGGTGGGAGAAACCAACGCTTACCTTACCGGTGT 240
DB 317 AAGCTCGGGGTATGCTTGTGGGTGGGAGAAACCAACGCTTACCTTACCGGTGT 376
QY 241 GTTATTGCTTACAGAGAGACATTTTCATGGGCTTACAAAGACAAATTCACAGCAAT 300
DB 377 GTTATTGCTTACAGAGAGAGATCTTATGGGCTTACAAAGACAAATTCACAGCAAT 436
```

```
QY 301 TACTTGGCTCCCTCAGCCCAAGCTCAGCTATTGAGTTGATTCCTTCAGGGAGAGAA 360
DB 437 TACTAGGCTCCCTCAGCCCAAGCTTCACATCAAGTGAAGCAAGTCAACAGGGAGAG 496
QY 361 ATCTATGCCCCATTCAGAGAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGG 420
DB 497 ATATATGGCCGCTTCAAGAGCTCTCTACCATGTGTGAACCCAGAGCAAGCTTGTGTTGGA 556
QY 421 GGATGGGATATGAGCAACATGAACTGCTGTATGTCATGAGCCAGGGCAAGAGTGTGTAAC 480
DB 557 GCGTGGAGATGAGCAGATGAATGCACTGGCAGATGCCATGACCAAGGCTGTGAGC 616
QY 481 ATGCAATTTGACAGAGAGTGAAGGCTTACATGATGATTCAGTTCCTCCCGGAAATC 540
DB 617 ATTTGACCTGCAAGAGAGCTCAGGCTTCATAGAGTCCATGCTGCACTTCCCGGTGTC 676
QY 541 TAAGACCCCGGATTTTCAATGCTGCCAACAAGAGAGCGTCCCAACAAGTGAATTAAGGCG 600
DB 677 TATGATCCCGGACTTTCATGCGCGCTAACAGAGGCTCTCGTGCAACAAATGTCATCAAGGCG 736
QY 601 ACAAGCAAGAGCAAGTTCAGCAATTCATCAAGACATTAAGCGCTTTAAGAGCAACC 660
DB 737 ACCAAGAAAGAACAGGTGAGAGATCATCAAAAGATATCAGGAGTTTAAAGAGAAAGAC 796
QY 661 AAAAGTGAACAAGGTGTTCTCTGTGACCTGCAACACAGAGAGGTATAGCAATTTGTT 720
DB 797 AAAAGTGAACAAGGTGTTCTGTGACCTGCAACACAGAGAGGTATAGCAATTTGTT 856
QY 721 GTAGGCTTAAATGACCAATGAGAAATCTCTGTGCTGTGTGACCAAGAAATGAGGCTGAG 780
DB 857 GCTGGCTCAACGACCAATGAGAAATCTGTGGCACTGTGTGACCAAGAAATGAGGCGGAG 916
QY 781 ATTTCTCTTCAACCTGATGCAATGCTGCTGTGATGGAAGAAATCTTCTTCAATTAAT 840
DB 917 ATCTGCGCAATCAACATATAGCAATGCTGTGTGACGAGAGGCTGTGCTTCAATTAAT 976
QY 841 GGAAGGCTTCAGAACTTTTGTACCAAGGCTGATGATCTTGCCATCCGAGAAACACT 900
DB 977 GGAAGGCTTCAGAACTTTTGTGCTGTGCTGTGATGATCTTGCCATCCGAGAAACACT 1036
QY 901 TTGATTTGTGAGATGATCTTCAAGAGTGTGACCAAGAAATGATCTGTGTTGTTGAT 960
DB 1037 CTGATGCGGTGTGACGACCTTCAAGAGTGTGACCAAGAAATGATCTGTGTTGTTGAT 1096
QY 961 TTTCTTGTGGGGCTGTGATCAAGCCAAATCTATAGTTAGTTCAACCATCTGGGAAC 1020
DB 1097 TTTCTTGTGGGGCTGTGATCAAGCCAAATCTGTGATGATGACCAACCTGGGAAC 1156
QY 1021 AATGATGATGATGATCTCTCGGCTCCAAACCTTCGCTCCAAAGAAATCTCAAGAGC 1080
DB 1157 AATGATGATGATGATCTCTCGGCTCCAAACCTTCGCTCCAAAGAAATCTCAAGAGC 1216
QY 1081 AATGATGATGATGATCTCTCGGCTCCAAACCTTCGCTCCAAAGAAATCTCAAGAGC 1140
DB 1217 AATGATGATGATGATCTCTCGGCTCCAAACCTTCGCTCCAAAGAAATCTCAAGAGC 1276
QY 1141 GACCATGTTGTTGTTATTAAGTATGCTCTTACATGAGGAGATGACAAAGAGCCATGAT 1200
DB 1277 GATCATGTTGTTGTTATTAAGTATGCTCTTACATGAGGAGATGACAAAGAGCCATGAT 1336
QY 1201 GAGTACCTTCAAGATATTTCAATGAGTGTGAAAGAACACCATGTTTGTGCAACACATGT 1260
DB 1337 GAGTACCTTCAAGATATTTCAATGAGTGTGAAAGAACACCATGTTTGTGCAACACATGT 1396
QY 1261 GAGGATTCCTTTTAACTGCTCTATTAATCTTGAATGCTGTCTTGTCTGAGCTGAGC 1320
DB 1397 GAGGATTCCTTTTAACTGCTCTATTAATCTTGAATGCTGTCTTGTCTGAGCTGAGC 1456
QY 1321 ACTAGAAATCAAGTTTAAAGTGAAGTGAAGGAAATTTCCATCTCAATCCACGATGTTGT 1380
DB 1457 ACCAGGATCAAGTTTAAAGTGAAGTGAAGGAAACGATCTCAATCTTCAATCCGATGTTGT 1516
```

```
QY 1361 ACCATTTCACTATCTGACCAAGGCTCTGTGTTCCACCGGGTACACCAAGTGTGAT 1440
DB 1517 ACCATTTCACTATCTGACCAAGGCTCTGTGTTCCACCGGGTACACCAAGTGTGAT 1576
QY 1441 GCATTGTCAAGCAGCGTGCATGCTGGAATAAATAAGAGGCTGTGTTGATTTGGCC 1500
DB 1577 GCTCTTGAAAGCAGAGGCGCATGCTGGAATAAATAAGAGGCTGTGTTGATTTGGCC 1636
QY 1501 CCAGAGAAATCAATGATCTCTGAGTACAAAGTGA 1533
DB 1637 CCAGAGAAATCAATGATCTCTGAGTACAAAGTGA 1669

RESULT 2
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197361
; GENERAL INFORMATION:
; APPLICANT: Martino-Cact, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytrate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10

Query Match 63.9%; Score 979.4; DB 3; Length 1931;
Best Local Similarity 77.4%; Pred. No. 5,8e-314;
Matches 1187; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 1 ATGTTTCAATGAGATTTTAAAGTGAAGTCTCTAATGTGAATGATACCGAGATGAGATT 60
DB 99 ATGTTTCAATGAGATTTTAAAGTGAAGTCTCTAATGTGAATGATACCGAGATGAGATT 158
QY 61 CAGTCCGTGTACAACTACGAAACACCGAACTTGTTCAGAGAAACAGAAATGCACTAT 120
DB 159 CAGTCCGTGTACAACTACGAAACACCGAACTTGTTCAGAGAAACAGAAATGCACTAT 218
QY 121 CAGTGGATTTGCAAAACCAATTCGTCAACTCAATTTAAACCAACCAATTTTCA 180
DB 219 CGCTGGGTGCTCGGCCCAAGTCCGTCAAGTACACTTCCGACCAAGACCGCTCC 278
QY 181 AAATTTGGGGGTATGCTGTGAGTGTGGGTGGAACCAACGCTTACCTCAACCGTGT 240
DB 279 AAGCTCGGGGTATGCTGTGAGTGTGGGTGGAACCAACGCTTACCTCAACCGTGT 338
QY 241 GTTATTTGCAAGAGAGACATTTCAATGAGGCTCAAGAGACAAAGTTCACCAAGCCAT 300
DB 339 GTTATTTGCAAGAGAGAGACATTTCAATGAGGCTCAAGAGACAAAGTTCACCAAGCCAT 398
QY 301 TACTTGGCTCCCTCAGCCCAAGCTCAGCTATTGAGTTGATTCCTTCAGGGAGAGAA 360
DB 399 TACTTGGCTCCCTCAGCCCAAGCTCAGCTATTGAGTTGATTCCTTCAGGGAGAGAA 458
QY 361 ATCTATGCCCCATTCAGAGAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGG 420
DB 459 ATCTATGCCCCATTCAGAGAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGG 518
```

QY 421 GATGGGATATCATCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTGTTGAC 480
DB 519 GGCCTGGGACATTTAGCAACATGAACCTGGCCGACTCCATGACCAAGGGCAAGGTGCTGAT 578
QY 481 ATCGATTTGCAAGAGCATTTAGAGCCCTTACATGAAATCCAATGTTCCATCTCCCGGAATC 540
DB 579 ATTGACCTGAGAGAGCATGCTCAGGCCCTTACATGAGATCCATGGTCCATCTCCCGGATC 638
QY 541 TACGACCCGGAATTCATTTGCTGCCAACCAAGAGAGCTGCCAACACATGATTTAAGGC 600
DB 639 TATGATCCGACCTTATCCGGGCTTAAACAGGGCTCTCCGCCCAACAGTGCATCAAGGGC 698
QY 601 ACAAGCAAGAGAGATGTCAGCAATCATCAAGACATCAAGCCGTTTAAAGAGCCACC 660
DB 699 ACCAAGAAAGAACAGGTGGAGAGCATCATCAAGATATCAAGGAGTTTAAAGAGAAAGAAC 758
QY 661 AAAGTGACAGAGGTGTTGCTCTGTGACCTGCCAACACAGAGAGATTAAGCAATTTGTT 720
DB 759 AAAGTGACAGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 818
QY 721 GATGGCTTAATGACACATGAGAAATCTTTGGCTGCTGTGAGACAGAAATGAGCTGAG 780
DB 819 GCTGGTCTCAACGACAGATGAGAAATCTAATGTCATCTGTGGAACAAGAGCGGAG 878
QY 781 ATTTCTCTTCCACCTTGTATGACATGCTGCTGTGATGAGAAATGTTCTTCTTCAATTA 840
DB 879 GTATCAACATCAACATATATGACATGCTGCTGTGATGAGAAAGGGTGCCTTATCAAT 938
QY 841 GGAGGCTCTAGAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCCGAGGAACACT 900
DB 939 GGGAGGCCCCAGAACCTTTGCTGCTGGCTGATTTGATCTTCTTATTAATAAACAACCTGC 998
QY 901 TTGATTTGAGAGATGACTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGTTGAT 960
DB 999 TTGATTTGAGTGTGACACTTCAAGAGTGTGACACCAAGATGAATCTGTGTTGTTGAT 1058
QY 961 TTTCTTTGGGGGCTGTATCAAGCCAACTCTATGTTAGTTACAAACCTCTGGGAAAC 1020
DB 1059 TTCTTTGTTGGTGTGTAATAAAGCCCACTCAATGTGTGATCAACCACTTGGGAAAC 1118
QY 1021 AATGATGATATGATCTCTCGGCTCCACAAACCTTCCGCTCAAGGAAATCTTCCAGAGC 1080
DB 1119 AAGAGTGGATGACCTGTCTGCTCCCTCAAAACATTCAGGTCCAGAGAGATCTTCCAGAGC 1178
QY 1081 AAGCTTTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAACCTCC 1140
DB 1179 AAGCTGTGATGACATGCTGTGACCAATGCCATCTCTATGAGCTGTGTAACCTCC 1238
QY 1141 GACCATGTTGTTTATTAAGATGTCCTTACATGAGGGATAGCAAGAGCCATGAT 1200
DB 1239 GATCATGTGTGTGTCAAGATATGTCCTGATGAGGAGACAGCAAGGGCTATGAGAC 1298
QY 1201 GAGTACACTTCAAGATATTCATGAGGTGAGAAAGAACCACTTGTTTTGGCAACACATGT 1260
DB 1299 GAGTACACTTCAAGATATTCATGAGGGGAGAGAAACCAATGCTGTGCAACACCTGT 1358
QY 1261 GAGGATTCCTTTTAACTGCTCTCTATTAATCTTGAACCTGTCTTCTGTGAGCTGAC 1320
DB 1359 GAGGACTCCTCTCCGCCCACTATCATCTGATCTGTGCTCTTGTGCTGAGCTCAGC 1418
QY 1321 ACTAGAATCAGTTTAACTGAAATGAGGAGAAATTCATCTCAATCCACATGTTGCT 1380
DB 1419 ACCAGGATCAGTGAAGCTGAGGAGAGAGCAAAATTCATCTTCCACCCGAGTGGCC 1478
QY 1381 ACCATTTCTAGCTATGACCAAGGCTCTCTGTGTTCCACCGGATCACCAAGTGTGAAT 1440
DB 1479 ACCATTTCTAGCTATGACCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1538
QY 1441 GCATTTCTAAAGAGCGTGCATGCTGAGAAACATTAAGAGGCTTGTGTGATTTGAGCC 1500
DB 1539 GCTGTGGCCAGAGAGGCGATGCTGAGAAACATCATGAGGCGTGTGAGGCTGTGCC 1598

QY 1501 CCAGAGATTAACATGATCTCGATGACATGTA 1533
DB 1599 CCAGAGAACACATGATCTTGGATCAAGTGA 1631

RESULT 3
US-09-677-064-10
; Sequence 10, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytrate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-10

Query Match 63.9%; Score 979.4; DB 3; Length 1931;
Best Local Similarity 77.4%; Pred. No. 5.8e-314;
Matches 1187; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 1 ATGTTTCATGAGAAATTTAAGGTAGAGAGTCTTATGTAATGTAACCGAGACTGAGATT 60
DB 99 ATGTTTCATGAGAGAGTCTCCGCTCGAGAGAGCCCAAGTGTGATGAGGCGGATGAGATC 158
QY 61 CAGTCCGTGTACATACAGAAACCAAGCACTTTTTCACAGAAACAGAAATGGCCATT 120
DB 159 GAGTGGAGTACCGGTACACAGAGAGCTGTGATACAGAGGCAAGAGAGAGAGAGAG 218
QY 121 CAGTGGATTTGCAAAACCCCAATCCGTCACATCAATTTAAACCAACACCAATGTTCA 180
DB 219 CCGTGGTGTGTCGCCCAAGTCCGTCAAGTCAACTTCCGACCAAGACCCCGTCCCG 278
QY 181 AAATGGGGGTATGCTGTGGGTTGGGGTGGAAACACGCTCTTACCTCAACCGGTGT 240
DB 279 AAGCTCGGGGTATGCTGTGGGGTGGGGAGGCAACACGAGGTTCACGCTGACGCGTGG 338
QY 241 GTTATTTCTAAGAGAGAGACATTTTCATGAGGCTTACAGAGAGAAAGATTTCAAGCCAT 300
DB 339 GTTCATTTCCAAAGAGAGAGATCTCATGAGGAGACCAAGAGAGAGAGAGAGAGAG 398
QY 301 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATTGAGTTGATCTTCCAGGAGAGAGAA 360
DB 399 TACTAAGGCTCCCTCACACGCTCTCAACATCAAGAGTCCGAGCTACACAGGAGAGAG 458
QY 361 ATCTATGCCCATTCAGAGAGTGTCTTCCATGTTAATCTGTGACGACATGTGTTGGG 420
DB 459 ATCTATGCCCATTCAGAGAGCTCTCTTCCATGTTAATCTGTGACGACATGTGTTGGG 518
QY 421 GGAATGGATTAATGAG 480
DB 519 GGCCTGGGACATTTAGCAACATGAACTGGCCGATCTCATGACCAAGGAGAGAGAGAG 578
QY 481 ATGATTTGACAG 540
DB 579 ATTGACCTGAG 638


```
QY 541 TACGACCCGATTTGATTTGCTGCAACCAAGAGAGCGTCCCAACAGTGAATTAAGGCG 600
Db 639 TATGATTCGGACTTCAATCGCGCTTAACAGGCGCTCTCGGCCCAAGTGTCTATCAAGGCG 698
QY 601 ACAAACAAGACCAAGTTCCAGCAAAATCATCAAAAGACATCAAGCGCTTTAAGAACCCACC 660
Db 699 ACCAAGAAAGAACAGGTGAGAGACATCATCAAGATATCAGGAGAGTTTAAGAGAAAGAAC 758
QY 661 AAAGTGGACAAAGTGTGTCTCTGTGAGATGCGCAACAGAGAGATATAGCAATTTGGT 720
Db 759 AAAGTGGACAAAGATGTGTGTGTGTGAGATGCGCAAAACATGAAAGATATAGCAATGTGTC 818
QY 721 GTAGGCGCTTAATGACCAACATGAGAGAAATCTTGTGCTGTGAGACAGAAATGAGGCTGAG 780
Db 819 GCTGCTCTCAACGACACGATGAGAGAAATCTAATGCGATCTGTGAGACAGAGAACAGCGAG 878
QY 781 ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGATGAGAAATGTTCTTTCAATTAAT 840
Db 879 GTATCAACCATCAACCTATATGCAATTCCTGTGTGATGAGAAAGGCTGCGTTCAATCAAT 938
QY 841 GGAAGCCCTCAAGACATTTTGTACAGAGGCTGATGATCTTGGCATCGCGAGAACACT 900
Db 939 GGGAGCCCCAGAACACCTTTGTGCTGTGCTGATGATCTTGTCTATMAAAACAACTGC 998
QY 901 TTGATTTGTGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTGTTGAT 960
Db 999 TTGATTTGTGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTGTTGAT 1058
QY 961 TTTCTTTGTGAGGAGTGTGATCAAGCCACATCTATAGTTACCAACATCTGGGAAATC 1020
Db 1059 TTTCTTTGTGAGGAGTGTGATCAAGCCACCTCAATGTGAGCTACCAACATCTGGGAAATC 1118
QY 1021 AATGATGATGATTAATCTCTCGGCTCAACAACTTCGCTCAAGAGAAATCTTCAAGAGC 1080
Db 1119 AAGGATGATGATTAATCTCTCGGCTCAACAACTTCGCTCAAGAGAAATCTTCAAGAGC 1178
QY 1081 AAGGTTGTGAGATGATGATCAAGCAATGCAATGCTCTATAGAGCTGTGTAACATCCC 1140
Db 1179 AAGGTTGTGATGATGATCAAGCAATGCTCTATAGAGCTGTGTAACATCCC 1238
QY 1141 GACCATGTTGTGATTAATGATGATGCTTCAAGAGGAGATGACAAAGAGAGCATGAT 1200
Db 1239 GATCATGTTGTGATTAATGATGATGCTTCAAGAGGAGATGACAAAGAGAGCATGAT 1298
QY 1201 GAGTACATCTCAAGATTAATGATGATGCTTCAAGAGGAGATGACAAAGAGAGCAT 1260
Db 1299 GAGTACATCTCAAGATTAATGATGATGCTTCAAGAGGAGATGACAAAGAGAGCAT 1358
QY 1261 GAGGATTTCCCTTTTACCTGCTCTATTAATCTTGAATGCTGCTTGTGAGCTGAGC 1320
Db 1359 GAGGATTTCCCTTTTACCTGCTCTATTAATCTTGAATGCTGCTTGTGAGCTGAGC 1418
QY 1321 ACTAAGATCAAGTTTAAGCTGAAATGAGGAGAAATTCACATCTATTCACCCAGTTGCT 1380
Db 1419 ACCAAGATCAAGTTTAAGCTGAAATGAGGAGAAATTCACATCTATTCACCCAGTTGCT 1478
QY 1381 ACCATTTCAAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGTTACACAGTGTGAAT 1440
Db 1479 ACCATTTCAAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGTTACACAGTGTGAAT 1538
QY 1441 GCATTTCAAGAGAGGCTGATGCTGAGAAACATATGAGAGGCTTGTGTGATTTGGCC 1500
Db 1539 GCCTTCCAGAGAGGCTGATGCTGAGAAACATATGAGAGGCTTGTGTGATTTGGCC 1598
QY 1501 CCAGAGATTAACATGATTTCTGAGATCAAGTGA 1533
Db 1599 CCAGAGATTAACATGATTTCTGAGATCAAGTGA 1631
```

RESULT 4
US-09-248-796A-3131
; Sequence 3131, Application US/09248796A

```
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248.796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 3131
/ LENGTH: 1578
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-3131
```

Query Match 25.1%; Score 384.8; DB 4; Length 1578;
Best Local Similarity 56.0%; Pred. No. 1.4e-116;
Matches 812; Conservative 0; Mismatches 602; Indels 36; Gaps 3;

```
QY 110 ATGCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 169
Db 113 ATGCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 172
QY 170 CCCATGTTCCAAATTTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 229
Db 173 TAAAGTCCCTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 232
QY 230 TCACCGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 289
Db 233 TGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 292
QY 290 AACAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 343
Db 293 TTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 352
QY 344 CCTTCAGGAGAGAGAAATCTATGCCCCCATGCAAGATCTGCTTCAATGTTAAATCTG 403
Db 353 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 412
QY 404 AGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 463
Db 413 ATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 472
QY 464 GGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 523
Db 473 GAGCTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 532
QY 524 TTCCACTCCCGGAATCTACAGCCCGGAATTTGATGATGATGATGATGATGATGATGAT 583
Db 533 AACCTTTGAATCAATCTACCTCCGATTTTATTTGATGATGATGATGATGATGATGAT 592
QY 584 ACAACGTGATTAAG-----GGACAAAGCAAGCAAGTTG 619
Db 593 ACAATGTTTAAACCAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 652
QY 620 ACGAATCTCAAGATCAAGGCTTTAAGAAAGCCCAAGATGATGATGATGATGATGATGAT 679
Db 653 AAAAAATCGAAAGATCAAGATTTCAAGGCCCAAAACAAATTAAGTAAAGTTATTA 712
QY 680 TCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739
Db 713 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 772
QY 740 TGAAGATCTCTTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 799
Db 773 CTGACAACTTGAATCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 832
QY 800 ATGCACTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 859
```

```
Db      833 TTGCGCTGCTTCTATCTTGGAAAAAGTTCCATATATTAAGTTCCACCAAAACACAT 892
Qy      860 TTGTACCAAGGCGTATGATCTTCCATCGCGAGAACTTTGATTTGGTGGAGTACT 919
Db      893 TTGTTCCCGTGTATTTAGTTAGTGAATAAATACACATCATTTGTTGGTGGAGTACT 952
Qy      920 TCAGAGGTGTAGACCAAAATGAATCTGTGTGGTTGATTTTCTTGTGGGGGGTGTGA 979
Db      953 TCAGAGCAGGTCAACAAATAATTAATCACTGTACTCAATTTCTTGTGGTGGTGTGA 1012
Qy      980 TCAGGCCAATCTATATAGTTAGTTACCAACATCTGGGAAACAAATGATGATGATCTCT 1039
Db      1013 TCAGAACATTTCTATTTGCTTCTTATATCACTTGGGTAACATGACGGTTACAAATTTAT 1072
Qy      1040 CGGCTTCACAAACCTTCCGCTCCAGAGAAATCTCCAGAGCAACGTTTGTGAGATATGG 1099
Db      1073 CATCACAAACCAATTTAGTATCTAAGAAATTTCCAAACATCTGTGTGTGAATATTA 1132
Qy      1100 TCAGACGAATGCGAT-----CCTGTATGAGCCTGTGTGAACATCCGACCATGTTGTTG 1153
Db      1133 TTGAATCCACGCAATTTATTTGTACACAGGAATCTGTGACAAAGTTGATCACTGTATCG 1192
Qy      1154 TTATTAAGTATGTGCTTATCTGTAGGGGATAGCAAGAGCCATGATGATCACTTCAAG 1213
Db      1193 TCATTAAATACCTTGCAGCGTGTGGTATTTCTAAGTTGCGATGATGATGATTAATCA 1252
Qy      1214 AGATATTCATGGGTGGAAGAACACATTTGTTGCAACACATGTGAGATTCCTTT 1273
Db      1253 AATTATGTGGGTGTGTCACAAATAATTAATTAATCAATGTTGTGAAGTTCATTCG 1312
Qy      1274 TAGCTGCTCTTATTTATCTTGAAGCTTGGCTCTTGTGAGCTGAGCAGCATAGAACAT 1333
Db      1313 TTGCTACACATTTGATTTATTTGACTTGTGTGTGCTACTGATTTGCACTAGAGTTCA 1372
Qy      1334 TTAAGCTGAAAAAGAGAAAAATTCACATTCACCCAGTTGCTACATTTCTGAGCT 1393
Db      1373 TCAGAGGCTCTGTATATCTGATTTATGATGATGATGATGATCTGTGTAGTTATTTGCT 1432
Qy      1394 ATCTGACCAAGGCTCTCTGTGTTCACCGGGTACACAGTGTGATGATGATGATCAAG 1453
Db      1433 ACTGCTCAAGGCTCATTTAGCAAGACAGGATTCAAACCTATCAACGATTAACAAAC 1492
Qy      1454 AGCGTCAATGCTGAAAAATTAATGAGGCTTGTGTGATTTGGCCCGACAGAAATACA 1513
Db      1493 AACGTCAACATTTAGTCAACCTTGTCAAGTGTGTGTGATTAACATTTGACAACTG 1552
Qy      1514 TGATTTCTCGA 1523
Db      1553 TAAGATTTGA 1562

RESULT 5
; US-09-397-787-34
; Sequence 34, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TUMOR THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-397-787-34
```

```
Query Match      22.8%; Score 349.4; DB 3; Length 1231;
Best Local Similarity 62.0%; Pred. No. 6.8e-105;
Matches 587; Conservative 0; Mismatches 356; Indels 4; Gaps 2;

Qy      584 ACAAGCTGATTTAAGGCAACAAAGCAAGCAAGTTTCAAGAAATCATCAAGATCAAG 643
Db      31 ACAACCTCATCCAGGCTCGGTGCGAGCACTGAGAGATATCCGACAGGACATCCGAG 90
Qy      644 CGTTTAAGAACCCACCAAAATGACACAGGTGTGTCTGTGTGACTGTGCAACAGAGA 703
Db      91 ACTTCGGCTGTAGCCCGGGGCTGACAAAGTATGTGTGTGTGTGACCGGGAAACAGAAC 150
Qy      704 GGTATAGCAATTTGGTTGAGGCTTTAATGACCAATGAGGAATCTTGTGGCTGTGG 763
Db      151 GCTTGTGTGAGGTGATTTCCAGGCTTCAACGACACAGCGAAACCTTGTGCGACATTTG 210
Qy      764 ACAGAAATGAGGCTGAGATTTCTCTTCCACCTTGTATGCAATGCTGTGTGTGATGAA 823
Db      211 AGCTGGTCTG---GAGGTGTGCTCTCCAGCTCTTGTGCGGTGCGACATCTGTGAGG 267
Qy      824 ATGTTCTTTTATTTAATGAGAACCTTCAAGACATTTTGTACAGGGCTGATTTG 883
Db      268 GCTGTGCTTCTCTCAATGAGGTCTCCGACAGAACCTTGTGCGGAGCTCTTGTGAGCTG 327
Qy      884 CCATGGCAGAGAACCTTTGATTTGGTGAATGATCTTCAAGAGTGTGACACCAAAATGA 943
Db      328 CGTGGACACCGGGTTTTTGTGGGCGAGATGACTTCAAGTCAAGCCGACCAAAATGA 987
Qy      944 AATCTGTGTGTGATTTTCTTGTGGGGCTGTGTATCAAGCAACATCTATAGTTAGTT 1003
Db      388 ACTCGGTGTGTGTGACTTCTCATTTGCTCGGCTCAAGACATGTCATGATGATTT 447
Qy      1004 ACAACCATCTGGGAAACAAATGATGATGA-TCTCTGGCTTCCACAAACCTTCCGCTCC 1062
Db      448 ACAACCATCTGGGAAACAAATGATGATGA-TCTCTGGCTTCCACAAACCTTCCGCTCC 507
Qy      1063 AAGGAATCTCCAAAGCAAGCTTGTGACGATGATGATGATGATGATGATGATGATGAT 1122
Db      508 AAGGAGGTGTCCAAAGCAAGCTTGTGACGATGATGATGATGATGATGATGATGATGAT 567
Qy      1123 GAGCTGTGTGAACATCCGACCATGTTGTATTTAAGTATGCTTTACCTAGGGGAT 1182
Db      568 AGCGCCGGGAAAGACCTGACACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 627
Qy      1183 ACAGAGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1242
Db      628 ACAGAGCCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Qy      1243 GTTTTGCACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1302
Db      688 GTGCTGACAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
Qy      1303 CTTTCTGCTGAGCTGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1362
Db      748 CTGCTGACCGAGCTGTGCTGAGGCTGAGCTTGTGATGATGATGATGATGATGATGATGAT 807
Qy      1363 TCATTTCAACCCAGTTGCTACATTTCTGAGCTATGATGATGATGATGATGATGATGATGAT 1422
Db      808 ACCTTCAACCCGCTGTGCTGCTGAGCTTCTTCAAGGGCCCACTAGTCCGCC 867
Qy      1423 GGTACACAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1482
Db      868 GGCAGCCCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
Qy      1483 GCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1529
Db      928 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 974

RESULT 6
; US-09-734-237B-74
; Sequence 74, Application US/09734237B
; Patent No. 6818752
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bui, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 74
/ LENGTH: 1605
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic gene derived from Saccharomyces cerevisiae myo-inositol
/ OTHER INFORMATION: -1-phosphate synthase, having numerous codons replaced with other
/ OTHER INFORMATION: s encoding the same amino acids to reduce free energy of folding,
/ OTHER INFORMATION: and a gly codon inserted after the initiating met codon
US-09-734-237B-74
```

```
Query Match      22.8%; Score 349; DB 4; Length 1605;
Best Local Similarity 55.8%; Pred. No. 1.1e-104;
Matches 779; Conservative 0; Mismatches 565; Indels 51; Gaps 4;
```

```
QY 180 AAAATTGGGGGATGATGCTTGTGGGTTGGGAAACAAGGCTCTACCCGCTGCG 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 195 AAGCGGTATATATGCTGATCGCTGGGTGTATCAACGGCTCTACTCTGTTGATC 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 240 TGTATTGCTTAA CAGAGAGACATTTCA TGGGCTACMAAGACAAGTTCA CAA 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 255 TGTTCGCAAA CAAACA CAGTAGAATTCAGACTAAGAGGTTAA CAGCCGA 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 300 TTACTTTGGCTCCCTCA CCAAGCTCA GCTATTCAGTTGG---ATCCTTCAGAGAGA 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 315 CTACTTGTGTTCTATGACTCA GGTGTTCTA CTGGAAGCTGGCAATGATGCTGAAGGTAA 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 357 GGAATATCTATGCCCCATTCAGAGTCTGCTCCAAATGTTAACTCGACGATTTGTT 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 375 CGACGTTAAGCTCGCTTA ACTCTGCTGCGATGATATCTCGAACGACTTGTTGT 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 417 TGGGGATGAGTATTCAGACA CATGAACCTGGCTGATGTCATGCGCAGGCAAGGTGT 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 435 TTCTGGTTGGATATCAACAGCGGATGTGTACGAGCAATGACAGCGTTCTCAGTTCT 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 477 TGAATCATGTTTGCAGAGCA GTTGAGGCTTACATGGAATCCATGTTCCACTCCCGG 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 495 GGAATATGATTCGCAACAGCGTCTGGAAGGCTAAGATGTCCTGTTAAAGCCACTGCGTC 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 537 AATCTACGACCGCGAATTTATGCTGCGCAACCAAGAGGAGCGTCCAAAC----- 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 555 CATCTACTACCCGATTTTATCGACGCTAACCAAGCGAACGCTTAACA CTGATCAA 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 -----AAGTATTAAAGGCA CAAAGCAAGACAAGTTCA CCAATATCAT 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 615 CCTGACGAAAGGGTAAGCTTA CTAACCGTGGTAAGTGA CTAACCTGCAAGGATCCG 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 630 CAAAGCATCAAGCGCTTAAAGAGCCA CAAAGTGA CAAAGGTGTTCTCTGTGAGC 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 675 TCGTGTATTCAGAACTTCAAGAGGAAACGCA CTGGA CAAAGTATGTA CTGTGGAGC 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 690 TGCCAA CAGAGAGATATGCAATTTGTTGAGGCTTAA TGACACATGAGAGATGT 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 735 TGTCTAACCTGAACGTTACGTAGAGATATCCCGGGTGAACGATACATGAGAAACCT 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 750 CTYGGCTGCTGAGACAGAAATGAGGCTGAGATTTCTCTCTTCCACTTGTATGCCATTGC 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 795 GGTGCAATTTATCAAGACGACCAAGAGAAATCGTTCGCTCA CCACTTGTGCTGCTGC 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 810 CTGTGTGATGAAATGTTCTTTCA TTAATGGAAGCCCTCAGACACTTTTGTACAGG 869
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
DB 855 ATCTATCTGGAAGCGGTACCGTACATCAAGGCTCTCCGAGAACACTTTGTA CCGGG 914
QY 870 GCTAATTGATCTTCCATTCGAGAGAA CATTGTA TTTGGTAGATGAC TTTCAA GAGTGG 929
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 915 TCTGTACAGCTGGCTGAAACGAAAGGTA CTTATCGCTGTGACGATCTGAATCTGG 974
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 930 TCAACCAAAATGAATCTGTGTTGGTTGATTTTCTTGTGGGGCTGTATCAAGCCAA C 989
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 975 CCAGACTAACTGAATCTGTATGTCGCA CAGTTCCTGGTTACCGCTGTATCAAA CCGGT 1034
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 990 ATCTATGTTAGTTACAA CCAATCTGGGAAA CAATGATGTA TGAATCTCGGGCTCCACA 1049
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1035 TTCTATCGCTCTTATTAACCA CCGTGGTTACAGACGCGCTTA CAACCTGTCTGCCGA 1094
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1050 AACCTTCGCTCCAAAGAAATCTCAAGAGCAAGCTTGTTCAGATATGATCAACAGCA 1109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1095 AAGTTCGTTCTTAAGAAATCTTAATCTCTGTATATGACGACATCATCGCTTCTAA 1154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1110 TGCCATCTCT-----CTATGAGCTGTGAACATCCGACCATGTTGTTTATTAAGTA 1163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1155 CGACATCCTGTACAA CCAAGCTGGTAAAGATGATCACTGTATCGTTATCAATA 1214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1164 TGTGCTTACGTAGGGGATACCAAGAGAGCCATGATGATCACTTCAGATATTCAT 1223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1215 CATGAACCGGTGTGATGATCTTAAGTTGCTATGACGAAATCTACTGTAACGTGATCT 1274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1224 GGGTGAAGAAACA CCAATGTTGTTGCA CACACATGTGAGGATTCCTTTAGCTGCTCC 1283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1275 GGGGGGTACAA CCGTATCTATTCACA CAGTTTGTGAACATCTGTCTGCTGCTACCC 1334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1284 TATTATCTTGA CTTGGTCTCTTCTTGTCTGAGCTGAGCACTGAATCCAGTTAA----- 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1335 GCTGATCATCA CACCGCTGTTATGACTGAATTCGTACCCGCTGATCTTCAAGAAAGT 1394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1338 -----AGCTGAAATGAGGGAATTCACATCATTTCCACCAGTTCGTAACATCT 1388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1395 TGACCCGGTTAAAGAAATGCTCGGCAATTCGAAACCTTACCCGGTTCGACTCTTCT 1454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1389 CAGCTATCTGA CCAAGGCTCTCTGTGTTCCA CCGGGTA CACAGTGTGAATGATGTC 1448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1455 GTCTTACTGGGTGAAGGTCCGGTGACGTG CAGGCTTCCACCCGGTTAACGGTGA 1514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1449 AAAGCAGCTGCATCTGGA AACAATATGAGGCTTGTGATTTGGCCCCAGAGAA 1508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1515 CAAACAGCGTACCGCTCTGGA AACTTCTCGTCTGTGATCGCCGCTCCAGAA 1574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1509 TAACTGATTCGCA 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1575 CGAAGTGGTTTCA 1589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 7

```
US-09-734-237B-72
/ Sequence 72, Application US/09734237B
/ Patent No. 6818752
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bui, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 72
/ LENGTH: 1602
/ TYPE: DNA
/ ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-72
```

Query Match 21.8%; Score 334.6; DB 4; Length 1602;
 Best Local Similarity 54.2%; Pred. No. 6,8e-100;
 Matches 847; Conservative 0; Mismatches 659; Indels 57; Gaps 6;

17 TTAAGTAGAGAGCTCTAATGTGAAGTACACCGAGACTGAGATTGAGTCGCGTGTACAACT 76
 35 TTAAGTAGATTACGACAAAGTGCACGTACAAAGACAAACGAGCTGTCTACCAATGACAGCT 94
 77 ACGAAACCAACCGAACTTTTACAGAGAACAGAAATGGCACCTTACGTGATGTCAAAC 136
 95 ACGAAAAATGCTG---TAGTTACGAAACAGCTAGTGGCCGCTTCGATGTAAACCCCACTG 151
 137 CCAATCCGCTCACTCAATTTAAACCAACCAACCATGTTCCAAATTTGGGGGTATGC 196
 152 TTCAAACATACGCTTTCAAACTTGAACCTTGAAGAAAGCCG---AATAACTAGAATTATGC 208
 197 TTGTGGGTTGGGGTGGAAACACGCGCTTACCTTCAACCGGTGTGTTATTTGTACAGAG 256
 209 TCATTGGGTAGGTGGCAACAATGGCTCCACTTAGTGGCTGGGTATGGCGAATAGC 268
 257 AGGACATTTTCATGGGCTACAAAGACAAAGATTCAACAGCAATTACTTTGGCTCCCTCA 316
 269 ACAATGTGAGATTCAAACTTAAGAAAGCGGTAAAGCAACCAACTTTCGCGCTCATGA 328
 317 CCCAAGCCTCAGCTATTGCGATTGATCCCTTC---AGGAGAGGAAATCTATGCCCCAT 373
 329 CTCATGTTCTTACTTGAACCTGGGTATCGATCCGAGGGGAATGACGTTATGCTCTT 388
 374 TCAGAGCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGGGGATGGGATATCA 433
 389 TTAACCTCTGTTGCCATGGTTAAGCCCAACGACTTTGTCGTCTGTGTGGGACATCA 448
 434 GCAACATGAACCTGCTGATGATCCATGCGCAAGGCAAAAGGTGTTGACATGATTTGCA 493
 449 ATAAACCAATCTATACGAAGCTATCAGAGATCAAGATTCGAATATGATCTCTCAAC 508
 494 AGAGTTGAGGCTTACATGGAATCGATGGTTCCATCCCGGAATCTACGACCGGAT 553
 509 AACGTTGAAGGCGAAGATGTCCTTGGTAGCCCTCTTCATTTACTACCTCGATT 568
 554 TCAATGCTGCCAACCAAGAGAGCGTGCNA-----CA 586
 569 TCATTCACAGTATCAAGATGAGAGACCAATACTGCATCAATTTGATGAAAAAGCA 628
 587 AGTGATTAAAGGCAAAAGCAAGCAAGTTCAGAAATCATCAAGACATCAAGCGCT 646
 629 ACGTAACCAAGAGGGTAAAGTGAGCCCATCTGCACGCAATCAGACCGCATATCAAAAT 688
 647 TTAAGGAAGCCACCAAGTGACAAAGTGTGTCCTGTGACATCGCCACACAGAGGT 706
 689 TCAAAGAAAAAGCGCCCTTGATAAGTAATCGTTCTTGGACTGCAAAATACAGAGGT 748
 707 ATAGCAATTTGGTTAGGCTTAAATGACACCATGAGAAATCTCTGGGCTGCTGAGCA 766
 749 ACGTAGAAGATATCTCTGCTGTTAATGACACCATGAAAACTCTTGCGAGTCTATTAA 808
 767 GAATAGAGCTAGATTTCTCTTCCACTTGTATGCCATGCTGTGTGATGAAAAATG 826
 809 ATGACCATGAAGAGATGCTCTCTTCCACGATCTTTCAGAGAGCATATCTTGGAAAGTG 868
 827 TTCCCTTTCAATTAAGAACCCCTCAGAACATTTTGTACAGAGGCGATTTGATCTTGCCA 886
 869 TCCCTATATTAATGTTGACCGCAAAATCTTTTGTCCCGCTTGTTGAGCTGGCTG 928
 887 TCGCAGAAACCTTTGATTTGTGAGATGACTTTCAGAGTGGTGCAGCAAAATGAAT 946
 929 AGCATAGGGTACATTCATTTGCGGAGAGCATCTCAAGTCGGGACAAACCAATTAAGT 988
 947 CTGTGTTGTTGATTTCTTGTGGGGCTGTGATCAAGCCCAATCTATATGTTAGTTACA 1006
 989 CTGTTCTGGCCCGAGTTCTTAGTGAAGCAAGGTATTAACCGGCTCATTTGATCTCTATA 1048
 1007 ACCATCTGGGAAACAAATGATGTGATATCTCTCGGCTCCACAAACCTTCGCTCCAAAG 1066

DB 1049 ACCATTTAGCAATATAGACGGTTATTAATCTATCTGCTCCAAAAACAATTTAGTCTAAAG 1108
 QY 1067 AAATCTCAAGAGCAACGTTGTTAGCATATGCTCAACGCAATGCTATCTCT----- 1120
 DB 1109 AGATTTCCAAAGTTCTGTCTATAGATGACATCATCGCTCTATATGATATCTTGTACATG 1168
 QY 1121 ATGAGCTGTGAACATCCCGACATGTTGTTGTTATTAAGTATGTCCTTAAGTGGG 1180
 DB 1169 ATAACTGGGTAAAAAGTTGACCACTGCTATGTCAATAAATATGAAGCCGCTGGGG 1228
 QY 1181 ATAGCAAGAGCCATGATGATGATCACTTCAAGATATTCATGGGTGAAAGAACCA 1240
 DB 1229 ACTCAAAAGTGGCAATGACGAGTATTAAGAGATGTTATGATGGGCTCAATACCGA 1288
 QY 1241 TTGTTTGCACAAACATGATGAGATTCCTTTTAAAGTCTCTTATATCTTGAATGG 1300
 DB 1289 TTTCCATTCACATGTTTTCGAAGATTTCTTACTGGCTACGCCCTTGATCATGATCTTT 1348
 QY 1301 TCCCTTCTGCTGAGCTGAGCATTAAGATCCAGTTTAAAGCT-----GAAA 1345
 DB 1349 TAGTCAATGACTGATTTGTATCAAGAGTGTCTTATAAAGGTGACCCAGTTAAAGAG 1408
 QY 1346 ATGAGGAAATTTCCACTCATTCACCCAGTGTCTACATTCAGCTATCTGACCAAG 1405
 DB 1409 ATGCTGGCAAAATTCGAAACTTTTATCAATTTTAACTTTGAGTTACTGTTAAAG 1468
 QY 1406 CTCCTCTGTTCCACCGGTTACACAGGTGTAATGATTTGCAAAAGCAGCTGCATGTC 1465
 DB 1469 CTCATTTACAGACAGAGATTTCAACCGGCTGAATGGCTTAAACAGAAAGACGGCT 1528
 QY 1466 TGGAAACATTAATGAGGCTTGTGTTGATTTGGCCCCAGAGATAATCATGATTTCTGAGT 1525
 DB 1529 TGAATAATTTTAAAGATTGTGATGATTCCTTCTCAAAACGAACTTAAGATTCGAAG 1588
 QY 1526 ACA 1528
 DB 1589 ACA 1591

RESULT 8
 US-09-118-442-15
 ; Sequence 15. Application US/09118442B
 ; Patent No. 6197561
 ; GENERAL INFORMATION:
 ; APPLICANT: Martino-Catc, Susan J.
 ; APPLICANT: Wang, Hongyu
 ; APPLICANT: Beach, Larry R.
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Bowen, Benjamin A.
 ; TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
 ; FILE REFERENCE: 0706
 ; CURRENT APPLICATION NUMBER: US/09/118,442B
 ; CURRENT FILING DATE: 1998-07-17
 ; EARLIER APPLICATION NUMBER: 60/055,446
 ; EARLIER FILING DATE: 1997-08-11
 ; EARLIER APPLICATION NUMBER: 60/055,526
 ; EARLIER FILING DATE: 1997-08-08
 ; EARLIER APPLICATION NUMBER: 60/053,944
 ; EARLIER FILING DATE: 1997-07-28
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FaSTSeq for Windows Version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 3546
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; US-09-118-442-15

Query Match 10.5%; Score 161.4; DB 3; Length 3546;
 Best Local Similarity 64.5%; Pred. No. 4e-42;
 Matches 305; Conservative 0; Mismatches 84; Indels 84; Gaps 1;

QY	979	ATCAAGCCAA	CATTTATAGTATG	TATTAACAACATCTGGGAAACATGATGATAACTC	1038
Db	2716	ATGAGGCCAC	CTCAATCTGTAGCT	TACACCACTTGGGAAACAAAGATGATGAACCTG	2775
QY	1039	TCGGCTCC	CAAACTTCCGCTCC	CAAGAAATCTCCAAAGACAAGTTGTTCAGATATG	1098
Db	2776	TCGCCCCCTT	AAACATTCAGGTCC	CAAGAGATCTCCAAAGCAACCTGTGTGATGACATG	2835
QY	1099	GTCAACAGCA	ATGCGATCCTAT	GAGCGTGTGATACATCCGACCATGTGTGTTATTT	1158
Db	2836	GTCTCGAGCA	ATGCGATCCTCTAT	GTAGAGCCGCGAGCATCCCGATCATGTCTTGTCTATC	2895
QY	1159	A-----	-----	-----	1159
Db	2896	AAGGTCTGTT	AGCTGATCTTT	CACCTCGTTAAAGTTGACATATGCAAGCAGATTTTACA	2955
QY	1160	-----	-----	AGTATGTGCTTACGTAGGGGATACGAAGAGCC	1194
Db	2956	TTGAACCTGT	CACCTCTTTTGTGTTAGT	GCAGATATGTGCCGATACGTGGAGACAGCAAGAGGGCT	3015
QY	1195	ATGATATAGTA	CATTCAGAGATATTA	TTCATGSGGTG3AAAGAAACAACATGTGTTTGACAAC	1254
Db	3016	ATGAGCAGAT	TAACCTCAGAGATCTT	CAATGSGGCGCAAGAAACAACATGTGTCTGCACAC	3075
QY	1255	ACATGTAGGA	TTCCCTTTTAGCTGCT	CTCTATTAATCTTGGACTTGTGTCCTTCTTGCTGAG	1314
Db	3076	ACCGTGAAGG	ACTCGCTCTTCGCGGAC	CTATATCTTGATCTGTG3TCTCTTGCGCTGAG	3135
QY	1315	CTGAGCACTGA	ATCCAGTTAAAGCTG	AAATATAGGGAATATCCATCTATTT	1367
Db	3136	CTGAGCACTGA	ATCCAGTTAAAGCTG	AAAGGATGATAGAGCCCTCCAAAGT	3188

```

RESULT 9
US-09-677-064-15
; Sequence 15, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytylate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677, 064
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-15

```

	Query Match	10.5%	Score 161.4	DB 3	Length 3546
	Best Local Similarity	64.5%	Pred. No. 4e-42		
	Matches 305	Conservative 0	Mismatches 84	Indels 84	Gaps 1
Oy	979	ATCAAGCCAATCTAGTATTGATACACACTGGAAAAAATGATGTATGAATTC	1038		
Db	2716	ATGCAGGCCAACTTAATCGTAGTGACACACCATTGGAAAACAAGATGGATGAACCTG	2775		
Oy	1039	TCCGCTCCACAAACCTTCGCTCCAGAAGAAATTCCAAAGACCAACGTTGTGACGATATG	1098		
Db	2776	TCTCCCTTCAACCATTCAGTCCAGGAGATCTCCAAAGACCAACGTTGTGATGAATG	2835		

QY	1099	GTCAAAGCAATGGCAATCCCTCTTATGAGCTGGGGAACATCCGACCATGTGTTGTATT	1158
Db	2836	GTCTCGAGCAATGCCATCTCTATGAGCCGGGAGCATCCCATCATGTCTTGTTCATC	2895
QY	1159	A-----	1159
Db	2896	AAAGTCTGTAGCTGATCTTCACTCTGTTAAAGTTGACATATGCAAGGCAATTACA	2955
QY	1160	-----AGTATGTGCTTTACGTAGGGGATAGCAAGAGGCC	1194
Db	2956	TTGAAACTTGTCACTCTTTTGTGACATGATGTGCCGATCTGGGAGACACAGAGAGGCT	3015
QY	1195	ATGATATAGTACACTTTAGAGATATTCAATGGTGGGAAGAAACCACTGTTTGGACAAC	1254
Db	3016	ATGAGCAGATACACTTAGAGATCTTCATATGGGGGCAAGAAACCAATCGTGTGCAAC	3075
QY	1255	ACATGTAGAGTTCCTCTTTAGTGTCTCAATATTATCTTGACATTGGTCTTCTTGCTGAG	1314
Db	3076	ACCGTATAGGACTCGCTCTCTGGCCGACCTATCATCTTTGATCTGGTGTCTTGGCTGAG	3135
QY	1315	CTGAGCACTAGAAATCCAGTTTAAAGCTGAAATAGAGGAAATTTCCACTATT	1367
Db	3136	CTCAGCACCAAGATCCACGCTGAAGCTGAGGAGAGAGTTAAGAGCCCCCAAGT	3188

```

; RESULT 10
; US-09-118-442-14
; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
; US-09-118-442-14

```

Query Match	10.4%;	Score 159;	DB 3;	Length 3546;
Best Local Similarity	64.3%;	Pred. No. 2.5e-41;		
Matches	304;	Conservative	0;	Mismatches 85; Indels 84; Gaps 1;
OY	979	ATCAAGCAATCTATAGTATGATTAACAACCTGGGAAACATGATGATGAATCTC	1038	
Db	2716	ATGACGCCACCTCAATCTGAGCTTACACCACTTGGGAAACACGATGGCATGACCTG	2775	
OY	1039	TCGGCTCCACAAACCTTCGCTCCACAGAAATCTCCAAAGACAGTGTGTGACGATATG	1098	
Db	2776	TCTCCCTTCAAAACATTCAGGTCCCAAGAGAAATCTCCAAAGACAGCGGTGTGATGACATG	2835	
OY	1099	GTCAAGCAATGCGCATCTCTATGAGCCGCGGTAAATCATCCGACCATGTGTGTATTT	1158	
Db	2836	GTCTCAGCAATGCGCATCTCTATGAGCCGCGGAGATCCGATCATGTGCTGTGCATC	2895	
OY	1159	A-----	1159	

```

Db      2896 AAGGCTGTAGTGAATCTTTGACCTCGTTAAAGTTGACATATGCAAGGACAGATTTACA 2955
Qy      1160 -----AGTATGTCCTTTAGCTAGGGATACCAAGAGGCC 1194
Db      2956 TTGAAACTTGTCACCTCTTTGTTGTCAGTATGTCCTACCTGCGAAGACAGAGAGGCGCT 3015
Qy      1195 ATGGATGATGACCTTCAGAGATATTCATGGGTGGAAGAACCACTTTGTTGCAAC 1254
Db      3016 ATGGAGGATGACCTTCAGAGATCTTCATGGGCGGCAAGAACCACTGCTGTCACAC 3075
Qy      1255 ACATGTAGAGATTCCTTTTATGCTGCTCTATATCTTGTGACTTGTCTTCTGCTGAG 1314
Db      3076 ACTGTAGAGACTCGCTCTCGCGCACCATCATCTGATGTGGTGTCTTGGCTGAG 3135
Qy      1315 CTGAGCACTAGATTCAGATTTAAAGCTGAAATGAGGAGAAATTCACATCTATT 1367
Db      3136 CTCAGCACCAAGATTCAGCTGAAAGCTGAGGAGGGGTAAAGCCCCCAAGT 3188

```

RESULT 11
US-09-677-064-14
Sequence 14, Application US/09677064
Patent No. 6291224

```

GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
FILE REFERENCE: 0706D
CURRENT APPLICATION NUMBER: US/09/677,064
CURRENT FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/055,446
PRIOR FILING DATE: 1997-08-11
PRIOR APPLICATION NUMBER: 60/055,526
PRIOR FILING DATE: 1997-08-08
PRIOR APPLICATION NUMBER: 60/053,944
PRIOR FILING DATE: 1997-07-28
PRIOR APPLICATION NUMBER: 09/118,442
PRIOR FILING DATE: 1998-07-17
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 3546
TYPE: DNA
ORGANISM: Zea mays
US-09-677-064-14

```

Query Match 10.4%, Score 159, DB 3, Length 3546,
Best Local Similarity 64.3%, Pred. No. 2.5e-41,
Matches 304, Conservative 0, Mismatches 85, Indels 84, Gaps 1,

```

Qy      979 ATCAAGCCCAACATCTTATGTTACCAACATCTGGGAAACAATGATGATGATCTC 1038
Db      2716 ATGAGGCCACCTCATCTGAGCTACCAACCTTGGGAAACAATGATGATGATCTC 2775
Qy      1039 TCGGCTCCACAACTTCGCTCCAGAAATCTCCAGAGCAAGCTTTGATGATGATG 1098
Db      2776 TCTGCCCTTCAAAACATTCAGGTCCAGAGATCTCCAGAGCAAGCTTTGATGATG 2835
Qy      1099 GTCAACAGAAATGCTCTTATGAGCTGTGGAACATCCGACCATTTGTTGTTATT 1158
Db      2836 GTCTGAGCAATGCTCTTATGAGCTGTGGAACATCCGACCATTTGTTGTTATT 2895
Qy      1159 A----- 1159
Db      2896 AAGGTCTGTAGTGAATCTTTGACCTCGTTAAAGTTGACATATGCAAGGACAGATTTACA 2955
Qy      1160 -----AGTATGTCCTTTAGCTAGGGATACCAAGAGGCC 1194
Db      2956 TTGAAACTTGTCACCTCTTTGTTGTCAGTATGTCCTACCTGCGAAGACAGAGAGGCGCT 3015
Qy      1195 ATGGATGATGACCTTCAGAGATATTCATGGGTGGAAGAACCACTTTGTTGCAAC 1254

```

```

Db      3016 ATGAGGATGACCTTCAGAGATCTTCATGGGCGGCAAGAACCACTGCTGTCACAC 3075
Qy      1255 ACATGTAGAGATTCCTTTTATGCTGCTCTATATCTTGTGACTTGTCTTCTGCTGAG 1314
Db      3076 ACTGTAGAGACTCGCTCTCGCGCACCATCATCTGATGTGGTGTCTTGGCTGAG 3135
Qy      1315 CTGAGCACTAGATTCAGATTTAAAGCTGAAATGAGGAGAAATTCACATCTATT 1367
Db      3136 CTCAGCACCAAGATTCAGCTGAAAGCTGAGGAGGGGTAAAGCCCCCAAGT 3188

```

RESULT 12
US-09-949-016-12608
Sequence 12608, Application US/09949016
Patent No. 6812339

```

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12608
LENGTH: 77626
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(77626)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12608

```

Query Match 6.0%, Score 92.6, DB 4, Length 77626,
Best Local Similarity 73.0%, Pred. No. 2.5e-18,
Matches 119, Conservative 0, Mismatches 44, Indels 0, Gaps 0,

```

Qy      983 ACCCAACATCTTATGTTACCAACATCTGGGAAACAATGATGATGATCTCTCG 1042
Db      34941 AGACCATGTCATCTGAGTTACCAACCTTGGGCAACAGATGGGAGAACTATCG 35000
Qy      1043 CTCACAAACCTTCGCTCCAGAGAAATCTCCAGAGCAAGCTTTGATGATGATG 1102
Db      35001 CGCCATTCGATTCGCTCTTAAGAGGTGTCCAGAGCAAGCTTTGATGATGATG 35060
Qy      1103 ACAGGAATGCCATCTCTATGAGCTGTGGAACATCCGAGCA 1145
Db      35061 AAGGAAACCAAGTGTCTATACGCCCGGCAAGAGCTTACCA 35103

```

RESULT 13
US-09-313-294A-4684
Sequence 4684, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 4684

LENGTH: 294
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6476212 700348858H1
NAME/KEY: unsure
LOCATION: 185, 272, 275, 281, 283, 288
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4684

Query Match 4.2%; Score 63.8; DB 4; Length 294;
Best Local Similarity 65.7%; Pred. No. 1.8e-10;
Matches 92; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 AAGTTCATCGAATTTTAAAGTAGAGTCTTAATGTGAAGTACACGAGCTGAGATT 60
DB 125 AATGTCATCGAAGCTTCGCGCTCGAGACCCCAAGTCCGTTACGCGCCGAGAGATC 184
QY 61 CAGTCCGTGTACAACAGCAACCAACGAACTTGTTCAGAGAACAGAAATGGACCTTAT 120
DB 185 NCTCGAGATACCGTAGACACAGACGAGCTGTGTCAAGAGCCAAAGGACGCGCTCC 244
QY 121 CAGTGGATTGTCAAAACCA 140
DB 245 CGCTGGTGTCTCGGCCCA 264

RESULT 14

US-09-902-540-5256
Sequence 5256, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 5256
LENGTH: 1335
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-5256

Query Match 2.9%; Score 44.2; DB 4; Length 1335;
Best Local Similarity 44.3%; Pred. No. 0.0017;
Matches 322; Conservative 0; Mismatches 393; Indels 12; Gaps 3;

QY 374 TCAGAAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGGGATGGGATATCA 433
DB 194 TCAGAAGCTGTGCTCTGCGGAGCTGAGAGACCTTCCGCGCTGGGACATCA 253
QY 434 GCAACATGAACCTGCGTGTGATGCCATGCGGCAAGGAGTTTGAATGATTTGACA 493
DB 254 TCCGTGAGAGCGATACGAAGTGGCGGTGCGCTCGGCGTGTCTCAAGCAAGCACCTG 313
QY 494 AGCAGTTGAGCGCTTACATGGAATTCATGTTTCACTCCCGGAATCTAGACCCGGATT 553
DB 314 AGAGGTGAAGCGCTTCTCTCAGAGCATCAAGCTGAAGGAGGCGGTGCAAGCTTGAAT 373
QY 554 TCATTGCTGCAACCAAGAGAGCGTGCACCAACCTGTATTAAAGGCAAAAGCAAGAC 613
DB 374 TGTGTGCGCGCATTCGAGGCAAC-----ACATCAAGGCAACCAAGAGCAACCGGAGA 427
QY 614 AAGTTCAAGAAATCATCAAGAGACATCAAGGCGTTTAAAGAGCCCAAAAGTGACAAAG 673
DB 428 GCATCGAAGCGCTGCGCCAGAGACATCCGGACTTCAAGAGAGAGCTCAACGCGACGCG 487

QY 674 TGTGTTCTCTGTGACTGCCAACACAGAGAGTATGCAATTTGTTAGAGCCTTAATG 733
DB 488 CCGTATGTGTGTGTCAGACAGCGCTGAGACCTTCGTCGCGCTGCCAGTCTTCA--- 544
QY 734 ACACCATGGAATATTTCTTGGCTGTGTGACAGAAATAGAGCTGAGATTTCTCTTCA 793
DB 545 AGACCTGTGCGCGCTTTCAGAAAGGCGCTGACAGAAAGCGCCGACATCAACCCACCG 604
QY 794 CTTGTATGCATTTGCTGTGTGATGAAATGTTCTTTTCATTATTAAGAGCCCTCAGA 853
DB 605 CGCTGTACCTTACCGCGCCATCAAGAGGCGGTGCTTGGCCAGCCACGCCCAAG 664
QY 854 ACACTTTTGTG---TACCAGGCTGATTTGATCTTGTCCATGCGAAGAACATTTGATG 910
DB 665 CCAGGTGAGACACGCGCGCGCTCCAGAGATGCGCAAGACGAGTCCGTGCGCGCGCG 724
QY 911 GAGATGACTTCAAGATGTGTACAGCAAAATGAAATCTGTGTGTTATTTTCTTGTG 970
DB 725 GCGCGAGCTTCAAGAGCGCGCCAGACGATGAAAGCGTATCTCGCGCGCTCAAG 784
QY 971 GCGGTGTATCAAGCCAAATCTATAGTTAATCAACCATCTGGGAAACATGATGTA 1030
DB 785 CCGCATCTGGGTGTGATGAGCTGTGTTCTCCACCAATCTGGGCAACCGCAGCGCG 844
QY 1031 TGAATCTCTCGGCTTCCAAACCTTCCGCTCCAGGAAATCTCCAAAGCAACGTTGTG 1090
DB 845 AGGTCTGAGAGACCCCGCGCTTCAAGGCGCAAGAGTCAACCAATGACGCTGTG 904
QY 1091 ACGATAT 1097
DB 905 ACACCAT 911

RESULT 15

US-09-902-540-1257/c
Sequence 1257, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1257
LENGTH: 34316
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-1257

Query Match 2.9%; Score 44.2; DB 4; Length 34316;
Best Local Similarity 44.3%; Pred. No. 0.017;
Matches 322; Conservative 0; Mismatches 393; Indels 12; Gaps 3;

QY 374 TCAGAAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGGGATGGGATATCA 433
DB 19705 TCAGAAGCTGTGCTCTGCGGAGCTGAGAGAGTGTGCTTGGCGCTGGGACATCA 19646
QY 434 GCAACATGAACCTGCGTGTGATGCCATGCGGCAAGGAGTTTGAATGATTTGACA 493
DB 19645 TCCGTGAGAGCGCATACGAAGTGGCGGTGCGCTCGGCGTGTCTCAAGCAAGCACCTG 19586
QY 494 AGCAGTTGAGCGCTTACATGGAATTCATGTTTCACTCCCGGAATCTAGACCCGGATT 553
DB 19585 AGAGGTGAAGCGCTTCTCTCAGAGCATCAAGCTGAAGGAGGCGGTGCAAGCTTGAAT 19526
QY 554 TCATTGCTGCAACCAAGAGAGCGTGCACCAACGATATTAAAGGCAAAAGCAAGAC 613


```
Db 19525 TCGTGGCCGCAATCGAGGCCAAC-----ACATCAAGGCCACCAAGACCGCACCGGAGA 19472
Qy 614 AAGTTAGCAAAATCATCAAGAGACATCAAGCCGTTTAAGGAAGCCACCAAGTAGACAAAG 673
Db 19471 GCATCCAGCGCGCCGACGACATCCGCACTTCAAGAGAGACTCAACGCGACCGCG 19412
Qy 674 TGGTTGTCCTGAGCATGSCAACAGAGAGATATAGCAATTGTTGTTAGGCTTAATG 733
Db 19411 CCGTATGTTGTTGTCAGAGACGTCGAGACCTTCCGTCCGCTGCCGAGTCTTCA--- 19355
Qy 734 AACACATGGAATCTCTTGGCTGTCGAGACAGAAATGAGGCTGAGATTTCTCTTCCA 793
Db 19354 AGACGCTGGCCGCTTCGAGAGCGCTGACAGAAACAGCCCGGACATCAACCCACCG 19295
Qy 794 CCTTATAGCCATTGCTGTGTGATGAGAAATGTTCTTTCATTAAAGAACCTTCAG 853
Db 19294 CGCTGACACCTACGCGGCATCAAGAGAGCGTGCCTTCGGAACGCCACGCCACAG 19235
Qy 854 AACCTTTG---TACGAGGCTGATGATCTTGCCATCGGAGGAACATTGATTGTTG 910
Db 19234 CCAAGCTGACACGCGCGGCTCCAGAGAGATGGCCAAGCAGAGTGGCGTGGCCGCG 19175
Qy 911 GAGATGACTTCAAGAGTGTCAAGACCAAAATGAAATCTGTGTGTTGATTTTCTTGTG 970
Db 19174 GCCGCACTTCAAGAGCGGCGACATGATGAGACGCTCATCGCGCCCGCTCAAG 19115
Qy 971 GGGCTGATCAAGCCAAACATCTATAGTTATACAACCATCTGGGAAACAAATGATGTA 1030
Db 19114 CCCGCAATGCTGGGCTGATGAGTGGTCTCCACCAACATCTGGGCAACCGGACGCG 19055
Qy 1031 TGAATCTTCGGCTCCAAACCTTCCGCTCCAGGAATCTCCAGAGCAAGCTTGTG 1090
Db 19054 AGGTGCTGACGACCCCGCGCTTCAAGGCCAAGAGTCAACCAAGTCGAGCGTGTG 18995
Qy 1091 ACGATAT 1097
Db 18994 ACACCAT 18988
```

Search completed: June 8, 2005, 03:03:36
Job time : 274.095 secs

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 22:54:36 ; Search time 5222.81 Seconds
(without alignments)
11172.621 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgttcacgcagaaattttaa.....tgattctcgcagracaaagta 1533

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	980.2	63.9	1740	9	CL961092 OeIFCC005
2	810.6	52.9	1628	3	BN827819 Arabidops
3	769.4	50.2	813	7	CK768601 Gm-r1030-
4	766.2	50.0	2360	1	AJ583520 AJ583520
5	698.8	45.6	988	7	CK277950 EST724028
6	651.4	42.5	914	7	CK275352 EST721430
7	648.4	42.3	936	7	CK279064 EST725142
8	636.2	41.5	865	7	CO085839 GR_Ea03E
9	633	41.3	906	7	CK320178 L2P08a02
10	628.2	41.0	968	7	CO113870 GR_Eb014
11	625.8	40.8	670	4	BG044525 BA029e07
12	624.4	40.7	860	7	CO072025 GR_Ea30P
13	623.6	40.7	842	7	CF211061
14	623.4	40.7	867	6	CB973407 CAB30002
15	620.4	40.5	891	7	CF212628 CGF10065
16	616	40.2	836	7	CF518404 CAP0007
17	610.2	39.8	618	2	AM279066 sg07402.Y
18	606	39.5	734	6	CA801137 saun01d09
19	597.2	39.0	746	6	CA918410 EST642557
20	593	38.7	594	4	BM523521 sam85a05
21	592	38.6	782	2	AM348136
22	591.2	38.6	860	7	CO082095 GR_Ea46I
23	590.2	38.5	762	2	BE660321 10B3 Gmax
24	588.6	38.4	707	2	BF005128 EST433626

c	25	587	38.3	754	4	BI968101	BI968101	GM830004A
c	26	586.8	38.3	834	7	CK933094	CK933094	CGF100434
c	27	585.4	38.2	978	4	BG388516	BG388516	GC01_01h1
c	28	583.2	38.0	831	7	CO132349	CO132349	GR_EB45C
c	29	582.2	38.0	672	2	AM508269	AM508269	8151h10.Y
c	30	581.2	37.9	883	6	CD438045	CD438045	EL01N0508
c	31	581	37.9	960	7	CK274458	CK274458	EST720536
c	32	580.4	37.9	622	2	AM397496	AM397496	8979e12.Y
c	33	575.4	37.5	830	7	CO107630	CO107630	GR_EB003
c	34	574.8	37.5	841	7	CO075472	CO075472	GR_Ea36C
c	35	573.6	37.4	816	6	CB975215	CB975215	CAB30006
c	36	573.2	37.4	663	6	CA900542	CA900542	PSC17211
c	37	573.2	37.4	891	7	CV290142	CV290142	acof01-10m
c	38	570.8	37.2	769	6	CA900545	CA900545	PSC19118
c	39	568.8	37.1	680	4	BG452447	BG452447	NF081D04L
c	40	563.6	36.8	670	2	BE191260	BE191260	8n94D09.Y
c	41	560.6	36.6	685	2	BF634972	BF634972	NF077A08D
c	42	560.4	36.6	648	1	AT794750	AT794750	8b68f07.Y
c	43	560.4	36.6	713	4	BI111216	BI111216	EST531296
c	44	557.6	36.4	585	4	BG239011	BG239011	8ab57d09
c	45	557	36.3	509	6	CB345158	CB345158	CA48BN000

ALIGNMENTS

RESULT 1	CL961092	1740 bp	DNA	linear	GSS 21-SEP-2004
LOCUS	OeIFCC005463	Oryza sativa	Express Library	Oryza sativa	(indica
DEFINITION	clutivar-group) genomic, genomic survey sequence.				
ACCESSION	CL961092				
VERSION	CL961092.1	GI:52376905			
KEYWORDS	GSS.				
SOURCE	Oryza sativa (indica clutivar-group)				
ORGANISM	Oryza sativa (indica clutivar-group)				
REFERENCE	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. M. and Wang, J.				
AUTHORS	1 (bases 1 to 1740)				
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis				
JOURNAL	Unpublished (2004)				
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.				
FEATURES	Location/Qualifiers				
source	1..1740 /organism="Oryza sativa (indica clutivar-group)" /mol_type="genomic DNA" /db_xref="taxon:39946" /clone_lib="Oryza sativa Expressed Library" /note="Oryza sativa exon trapped genomic sequences"				
ORIGIN					
Query Match	63.9%; Score 980.2; DB 9; Length 1740;				
Best Local Similarity	77.6%; Pred. No. 2,7e+285;				
Matches 1186; Conservative	0; Mismatches 343; Indels 0; Gaps 0;				
QY	1 ATGTTTCATCGAGATTTTAAGGTAGAGCTCTTAATGTGAGTACACCGAGCTAGATT 60				
DB	1 ATGTTTCATCGAGAGCTTCGCGTGGAGAGCCGCGACGTCGCGTACGCGCGCGAGATC 60				
QY	61 CAGTCCGTGTACAACTAGAAACACCGAAGCTTGTTCACGAGAAAGAGATGCACTAT 120				

```

Db      61 GAGTGGAGATACAGATACAGACAGAGAGCTGTGTGACAGAGACAGACAGAGCGCCCTCC 120
Qy      121 CAGTGGATTTGTAAACCCAAATCCGCTCACTACCAATTTAAACCAACCCCATATGTCCA 180
Db      121 CGCTGGGTGTGTCCGCCAAAGTCCGTCCGTCAACTTCGAGACCAACCAACCGTCCCTCC 180
Qy      181 AAATGGGGGTATGATCTTGTGGGTGTGGAGAAACAAGGCTTACCTTCACCGGTGAT 240
Db      181 AAGCTCGGGGTATGTCTGTGGGTGTGGGCGGCAACAAGGCTTCAACGTGTGAGCGGTGGG 240
Qy      241 GTTATTTGCTAACAGAGAGACATTTTCATGGGCTTACAAAGACAAAGATTCACAGCCAT 300
Db      241 GTCATTCGCAAGAGGAGGAAATTCATATGGGCGACCAAGAGCAAGGTGACAGCAACCAAC 300
Qy      301 TACTATTTGGTCTCCCTACCCAAAGCTCAGTATTCAGATTTGATTCCTCCAGGAGAGAA 360
Db      301 TACTATTTGGTCTCCCTACCCAAAGCTCAGTATTCAGATTTGATTCCTCCAGGAGAGAG 360
Qy      361 ATCTATGCCCCATTCAGAGAGTCTGCTTCCATATGTTAATCTGACGACATTTGTTGGG 420
Db      361 ATCTATGCCCCCTTCAAGAGCTCTCTGCCCATGTGTGAACCTTGATGACCTTGTGTGGG 420
Qy      421 GGATGGGATATACGACAAATGAACTTGCTGTATGCCATGAGGCGCAAGGCGAAAGTGTGAC 480
Db      421 GGCTGGGACATTAAGCAATGAACCTGCTGTATGATGACAGGCGCAAGGCTACTTGAC 480
Qy      481 ATCGATTTGGCAGAGAGATTTGAGGCGCTTACATGGAATCCATGTTCCACTCCCGGAATC 540
Db      481 ATTTGATCTGCAAGAGAGCTCAGACCTTTACATGATGATGATGATGATGATGATGATGAT 540
Qy      541 TAGACACCCGATTTTCAATGCTCTGCCAAACAAGAGAGCGTGCACCAACGCTGATTAAGGCG 600
Db      541 TATGACCCCGACTTCAATGCTCTGCCAAACAAGAGATCCCGGCGCAACAATGCTATCAAGGCG 600
Qy      601 ACAAAAGCAAGACCAAGTTGACGAAATCATCAAAAGACATCAAGGCGTTTAAGAGCCACC 660
Db      601 ACCAAGAGAGACCAAGTGTGAGAGATCATCAAGAGACATCAAGGAGTTCAAGAGAAAGAGC 660
Qy      661 AAGATGGACAAAGTGTGTCTCTGTGAGATGTCACACAGAGAGATGATGATGATGATGAT 720
Db      661 AAGATGGACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy      721 GTAGGCTTAAATGACACCAATGAGAGATCTTTGCTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db      721 GTTGGGCTAATGACACCAATGAGAGATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy      781 ATTTCTCTTCCACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db      781 ATATCAACATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy      841 GGAAGCCCTCAGAAACCTTTGTACAGGCGTGTGATGATGATGATGATGATGATGATGATGAT 900
Db      841 GGAAGCCCTCAGAAACCTTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Qy      901 TTGATTTGTGTGAGATGATCTTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Db      901 CTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Qy      961 TTTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db      961 TTTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Qy      1021 AATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db      1021 AATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Qy      1081 AATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db      1081 AATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Qy      1141 GACCATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200

```

```

Db      1141 GATCATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Qy      1201 GAGTACACTTTCAGAGATATTTTCATGAGTGTGAAGAACAACCATTTGTTTTCACAAACATATG 1260
Db      1201 GAGTACACTTTCAGAGATATTTTCATGAGTGTGAAGAACAACCATTTGTTTTCACAAACATATG 1260
Qy      1261 GAGATTCCTTTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db      1261 GAGGACTCATTTCTTGTCCGCCCATCATTTCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Qy      1321 ACTGAAATCCAGTTTAAAGCTGAAATAATGAGGAAATTCCTCATTCACCAAGTGTGT 1380
Db      1321 ACCAGATTCAGCTGAAAGCCGAGGAGAGAAATTCATTCCTTCATTCAGTGTGT 1380
Qy      1381 ACCATTCAGCTATTCAGCAAGGCTCTCTGTGTTCACCGGATACCAAGTGTGTAT 1440
Db      1381 ACCATTCAGCTATTCAGCAAGGCTCTCTGTGTTCACCGGATACCAAGTGTGTAT 1440
Qy      1441 GCATTGTCAAAAGCAGCGTGAATGCTGAAAAATATATGAGGCGTTGTGTGATTTGCTC 1500
Db      1441 GCCCTGGCAAAAGCAGAGGCAATGCTTGAGAAATCATGAGGCGCTGTGTGTGTGTGTGTGT 1500
Qy      1501 CCAGGAAATTAATATATTCCTGAGTACAA 1529
Db      1501 CCGGAAACCAATGATCTTGAGTACAA 1529

RESULT 2
CN50A4MK      1828 bp      mRNA      linear      HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA complete sequence from clone
GSLTPGH242E06 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
BX827819
BX827819.1 GI:42462602
HTC; GS.LT_cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1828)
REFERENCE
AUTHORS
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Querier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
JOURNAL
REFERENCE
AUTHORS
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
_length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
source
1..1828
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"

```

gene
/db_xref="taxon:3702"
/clone="GSLRGH242B6"
/issue_type="Hormone Treated Callus"
/plasmid="pCMVSPORT_6"
/complement(1..1828)
/gene="At4g39800"

ORIGIN

Query Match 52.9%; Score 810.6; DB 3; Length 1828;
Best Local Similarity 71.7%; Pred. No. 8.4e-234;
Matches 1108; Conservative 0; Mismatches 424; Indels 13; Gaps 3;

1 ATGTCATGAGAAATTTAAGTAGAGAGTCCCTAATGTGAAGTACCGAGCTGAGATT 60
58 ATTTTATTGAGAGCTTCAAAGTTGAGACCCGAACTGAAGTACAGAGAAATGAGATT 117
61 CAGTCGTGTACAGTACGAAACCAACGAACTTGTTCAGAGAA---CAGAAATGAGACC 117
118 CACTGTGTGATGATTAAGAGAGAGAGGTTGTTCAAGAGAAAGCCGTCAATGGCACT 177
118 TATCAGTGTGATGATTAAGAGAGAGAGGTTGTTCAAGAGAAAGCCGTCAATGGCACT 177
178 TACCAATGATGATGATTAAGAGAGAGAGGTTGTTCAAGAGAAAGCCGTCAATGGCACT 237
178 CCAAAATGGGGGTGATGCTGTGGGTGGGGTGGAAACAAGGCTCAACCGGT 237
238 CCAAAATGGGGGTGATGCTGTGGGTGGGGTGGAAACAAGGCTCAACCGGT 297
238 GGTGTTATGCTTAACAGAGAGACATTTGATGGGCTTACAAAGACAAAGATTAACAGACC 297
298 GGTGTTATGCTTAACAGAGAGACATTTGATGGGCTTACAAAGACAAAGATTAACAGACC 357
298 AATTACTTGGGCTCCCTCAACCAAGCTCAAGCTTATGAGTTGATCTTTCAGAGAGAG 357
358 AATTACTTGGGCTCCCTCAACCAAGCTCAAGCTTATGAGTTGATCTTTCAGAGAGAG 417
358 GAAATCTATGAGGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 417
418 GAGATCTATGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 477
418 GAGAGTGGGATATAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 537
478 GAGAGTGGGATATAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 537
478 GAGATCTATGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 537
538 GATATCTGACTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 597
538 ATCTGAGACCGGAGTTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 597
538 ATCTGAGACCGGAGTTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 657
538 GGCACAAAGCAGAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 657
658 GGTACCAAGAGAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 717
658 ACCAAAGTGGACAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 717
718 AATAAGTGGATTAAGGTTGTTCTCTGAGCGGCTTAACAGAGGTTATAGCATATG 777
718 GTTGTAGGCTTAAAGACCAATGAGAAATCTTGGCTGTGAGCAAGAAATGAGGCT 777
778 GTTGTAGGCTTAAAGACCAATGAGAAATCTTGGCTGTGAGCAAGAAATGAGGCT 837
778 GAGATCTCTCTTCAACATTTATGAGATTTGTTGTTCTTGAAGGAAATCCGTTTCACT 837
838 AATGAGAGCTTCAACATTTATGAGATTTGTTGTTCTTGAAGGAAATCCGTTTCACT 897
838 AATGAGAGCTTCAACATTTATGAGATTTGTTGTTCTTGAAGGAAATCCGTTTCACT 897
898 AATGAGAGCTTCAACATTTATGAGATTTGTTGTTCTTGAAGGAAATCCGTTTCACT 957
898 ACTTTGATTTGAGAGATCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAGCTT 957

958 GTTTGATCGGTGAGAGATTTTCAAGAGTGTCTCAACCAAGATGAATCTGTCTTGGTT 1017
958 GATTTTCTTGTGGGGCTGTGATTAAGCAACATCTATATGTTAGTTACCAACATCTGGGA 1017
1018 GATTTCTAGTTGGGCGCAGAAATCAAGCTTCAATGTGTAGCTACCAATCACTTGGGA 1077
1018 AACATGATGATTAATCTCTGGGCTCCAAACCTTCCGCTCAAGAAATCTTCAAG 1077
1078 AACCAAGATGATTAATCTCTGGGCTCCAAACCTTCCGCTCAAGAAATCTTCAAG 1137
1078 AGCAAGTTGTTGAGATATGTTCAAGCAAGATGCTCTTATAGAGCTGTGAAACAT 1137
1138 AGCAATGCTGTGAGAGATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1197
1138 CCGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1197
1198 CCGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1256
1198 GATGATGATCTTCAAGATATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
1257 AGGAGTTTAAAGAGAGAAATTAAGGATTAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1316
1258 TGTGAGATTCCTTGTGATGCTCTTATATCTTGAATGATGATGATGATGATGATGATGATGATGAT 1317
1317 GAGCTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376
1318 AGCAATGAT 1368
1377 GTGATTAAGAGATGAGCTGAGATCTCTCTTCAACATTTATGAGAAATGATCACTTTC 1436
1368 CACCAATGTTGTTGATCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1428
1437 CATCTGTGGCCACCATCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
1429 CCAAGTGTGAT 1488
1497 CCGATGATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556
1489 GTTGTGAT 1533
1557 GTTGTGAT 1601

RESULT 3
CK768601
LOCUS
DEFINITION
Gm-r1030-621 Gm-r1030 Glycine max cDNA clone Gm-r1030-621 5', mRNA
SEQUENCE
ACCESSION
CK768601
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
EST.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 813)
Periapurum, C.C., Li, L., Wurtele, E.S., Westgate, M.E. and
Nikolaou, B.J.
Molecular determination of soybean composition
Unpublished (2004)
Contact: Basil J. Nikolaou and Cyril C. Periapurum
Department of Biochemistry, Biophysics and Molecular Biology
Iowa State University
2210 Molecular Biology Building, Ames, IA 50011, USA
Tel: 515 294 9423
Fax: 515 294 0453
Email: dimma@iastate.edu
Individual base call and confidence value were assigned using the
phred software (http://www.phrap.org/). Overall sequence quality
assessment and vector trimming were conducted using the Lucy

software (<http://www.tigr.org/software/>).
This clone was originally generated by the Public Soybean EST
Project (<http://129.186.26.94/soybeanest.html>) /Shoemaker, R
(reshoe@iastate.edu).
This clone is available through: Biogenetic Services, 801 32nd Ave,
Brookings, SD 57006. For further information call 605-697-8500 or
contact info@biogeneticservices.com
Seq primer: T7-1 (5' AAT ACG ACT CAC TAT AG 3').
Location/Qualifiers

FEATURES

source

1..813
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1030-621"
/lab_host="DH10B"
/clone_lib="Gm-r1030"
/note="vector: pSPORT1, Site 1: SalI, Site 2: NotI, This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants
(individual seed fresh weight of 100-300mg). The library
was prepared using the life Technologies pSuperScript cDNA
library construction kit. Complementary DNA was
synthesized from mRNA using a poly(dT) sequence with a
NotI restriction site. SalI linkers adapters were ligated
to the blunt-ended cDNA fragments followed by NotI
digestion. The cDNA fragments were directionally cloned
into the NotI-SalI restriction site of the pSPORT1
vector. The ligated cDNA fragments were transformed into
E. coli Electromax DH10B host cells. This library was
constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note
that Gm-r1030 is a re-rack of Gm-cl007."

ORIGIN

Query Match 50.2%; Score 769.4; DB 7; Length 813;
Best Local Similarity 97.4%; Pred. No. 1.9e-221;
Matches 782; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

432 CAGCAACATGAACCTGGCTGATGCCATGGCCAGGCAAGTGTGACATGATTTGCA 491
11 CGGCAACATGAACCTGGCTGATGCCATGGCCAGGCAAGTGTGACATGATTTGCA 70
492 GAAGCAGTTGAGGCTTACATGGAATCCATGCTCCCGGAATCTACAGCCGGA 551
71 GAAGCAGTTGAGGCTTACATGGAATCCATGCTCCCGGAATCTACAGCCGGA 130
552 TTTCATGTGTCACCAAGAGAGAGGCTGCCAACAACGATTTAAGGCAACAAGCA 611
131 TTTCATGTGTCACCAAGAGAGAGGCTGCCAACAACGATTTAAGGCAACAAGCA 190
612 GCAAGTTCAGCAATCATCAAGAATCAAGGCGTTTAAAGAACCAACAAAGTGACNA 671
191 GCAAGTTCAGCAATCATCAAGAATCAAGGCGTTTAAAGAACCAACAAAGTGACNA 250
672 GGTGGTGTCTGTGAGACTGCCAACAAGAGATTAAGCAATTTGGTGGCCCTTAA 731
251 GGTGGTGTCTGTGAGACTGCCAACAAGAGATTAAGCAATTTGGTGGCCCTTAA 310
732 TGACACCATGAGAACTCTTGGCTGCTTGACAGAAATGAGCGTGAATTTCCCTTC 791
311 TGACACCATGAGAACTCTTGGCTGCTTGACAGAAATGAGCGTGAATTTCCCTTC 370
792 CACCTTGATGCAATGCGCTGTGTGATGGAATAATGTTCTTTCAATTAAGAACCCCTCA 851
371 CACCTTGATGCAATGCGCTGTGTGATGGAATAATGTTCTTTCAATTAAGAACCCCTCA 430
852 GAACACTTTTGTACACGAGGCTGATTTGATCTTGGCATCGGAGAAACATTTGATGGG 911
431 GAACACTTTTGTACACGAGGCTGATTTGATCTTGGCATCGGAGAAACATTTGATGGG 490
912 AAGATGACTCAAGAGTGTGACAGCCAAATGAAATCTGTGTGATTTCTTGTGGG 971
491 AAGATGACTCAAGAGTGTGACAGCCAAATGAAATCTGTGTGATTTCTTGTGGG 550

QY 972 GGCTGTATCAAGCCAAATCTATAGTTAGTACCAACCATTTGGGAAACAATGATGTTAT 1031
DB 551 GGCTGTATCAAGCCAAATCTATAGTTAGTACCAACCATTTGGGAAACAATGATGTTAT 610
QY 1032 GAATCTCTGGCTCCACAAACCTCCGCTCCAAAGAAATCTCCAAAGCAAGTGTGGA 1091
DB 611 GAATCTCTGGCTCCACAAACCTCCGCTCCAAAGAAATCTCCAAAGCAAGTGTGGA 670
QY 1092 CGATATGGTCAACAGCAATGCGCATCTATAGAGCTGTGTGAGATCCGACCATGTTGT 1151
DB 671 TGATATGTGTCAACAGCAATGCGCATCTATAGAGCTGTGTGAGATCCGACCATGTTGT 730
QY 1152 TGTATTTAGTATGTGCTTACGTAGAGGAGATGACAGAGCCATGATGATCTTC 1211
DB 731 TGTATTTAGTATGTGCTTACGTAGAGGAGATGACAGAGCCATGATGATCTTC 790
QY 1212 AGAGTATTTCTGGGTGGAAGA 1234
DB 791 AGAGTATTTCTGGGTGGAAGA 813

RESULT 4

AJ583520/c 2360 bp mRNA linear EST 08-JAN-2004
LOCUS AJ583520 PpProcl Physcomitrella patens cDNA clone pp001015066
DEFINITION similar to putative myoinositol-1-phosphate synthase, mRNA
sequence.

VERSION AJ583520 GI:40781590
KEYWORDS EST.

SOURCE

ORGANISM

Physcomitrella patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
1 (bases 1 to 2360)
Hohe,A., Egner,T., Lucht,J.M., Holterf,H., Reinhard,C., Schween,G.
and Reekl,R.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..2360
/organism="Physcomitrella patens"
/mol_type="mRNA"
/db_xref="taxon:3218"
/clone="pp001015066"
/tissue_type="protonema"
/clone_lib="PpProcl1"

ORIGIN

Query Match 50.0%; Score 766.2; DB 1; Length 2360;
Best Local Similarity 70.3%; Pred. No. 2.8e-220;
Matches 1088; Conservative 0; Mismatches 443; Indels 16; Gaps 4;

1 ATGTTCAATGAGAAATTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAAGTGAATT 60
DB 2150 ATGTTGTCGACAGCTTTCGTTGTAGAGTCTCTCATGTGGAATACACGAGACATGATC 2091
QY 61 CAGTCGGTGTCAACTAGAAACACGGAACCTGTCTCAGG-----AACGAAATGGC 114
DB 2090 AAGTGAAGCTATGATGACCAACACGAGTTCCTGCAAGACGCAAGAGCATGGC 2031
QY 115 ACCTATCAGTGGATTTGCAAAACCAATCCGTCAACTCAATTTAAACCAACCAACCAT 174
DB 2030 AAGTATGATGGGTAGCAAGCCCAAGTCCGTGATGCAATTTGCCACTCAGGCAAG 1971
QY 175 GTTCAAAATTTGGGGTATAGTCTTGTGGTGGGTGGAACAACAGGCTTACCTCAAC 234

Dp	1970	GTGCTTAAGCTGGGAATGATGCTCGTGGGTGGGCAACAGGGTCCACTCTCACG	1911
Oy	225	GGTGTTGTTATTTGCTTAACAGAGAGACATTTTCATGGGCTACAAAGA-CAAGATTCAACA	293
Dp	1910	GCTGTTATCTTCGCAACAAAGAAAGTATCTCAATGGGTGACCAAGAGATGGCGTTGACGA	1851
Oy	294	AGCCAAATTAATTGGCTCCCTCCACCCAGAGCTCAGACTATTGAGTTGGATCTTCC---	349
Dp	1850	CGCCAAATTACTCGGATCCCTCAACAAAGCTCATTGGACGGCGGTCTCTTTCACAG	1791
Oy	350	----AGGAGAGAGAAATCTATGCCCCCATTTCAAGAGTCTGCTTCCATGTTTAATCTGAC	405
Dp	1790	GTGGAAGAGAAATTCATGATGACTTTTAAAGAACATTTAAACCAATGTTCAACCCCAAC	1731
Oy	406	GACATTTGTTTGGGGAGATGAGATATGACAACATGAACCTGGCTGATGCCATGGCCAG	465
Dp	1730	GATCTCGTGAATCGGCGTTGGGATATCTTCCAAATGAATCTTGCGGATGGCATGGATGCT	1671
Oy	466	GCAAAAGTGTTTGACATGATTTTGGAGAAGCAGTTGAGGCTTACATGGAATCCATGTT	525
Dp	1670	GCCCCGCTGTCGAGATGATCTAACAGAGCACTTCCTCTCTCATGACAGATATGTTA	1611
Oy	526	CCACTCCCCGGAATCTAGACCCGGATTTCATGTGTCGAACCAAGAGAGCGTGGCAAC	585
Dp	1610	CCCATGCTGGTATTTACGACCCCGACTTCATCGCTGTAAACAGAGAGCTGTGGCAAC	1551
Oy	586	AACGTGATTAAGGCAACAAAGCAAGCAAGTTGACGAAATCATCAAGACATCAAGCG	645
Dp	1550	AATCTCATCAAGGGGTTCCAAAGAGAGCAATGAGCGGATTTATCCAGGACATACGGAAC	1491
Oy	646	TTTAAGGAAGCAACCAAGTGGACAAAGTGTGTTGCTGTGACTGCCAACAAGAGG	705
Dp	1490	TTCAAGGGGCAACCAAGTGTGACAAAGTATGTGATCTGTGACCGCTTAAACAGAGCGT	1431
Oy	706	TATAGCAATTTGGTGTGAGGCTTAATGAC-CCATGAGAAATCTTGGCTGCTGTCGA	764
Dp	1430	TACAGCAGTGAATGTGTGATGATATACCCAAAGAGATCTGTTGATTTCTTGA	1371
Oy	765	CAGAATGAGCTGAGATTTCTCCTCCACCTTGATCCACTTGCCTGTGATGAGAA	824
Dp	1370	GAAAGGAAATGAAATCTGCGCGTGCACCTTGTAGCTCTCGCTTGCAATCCAGAGCA	1311
Oy	825	TGTTCTTTCAATTAAGAGCCCTCAGAACATTTGTATCCAGGGCTGATGTACTTGC	884
Dp	1310	GGTTCCTTCATCAACGGCAGTCCCCAGAACACATTCGCTCCGTGAGTAATGAATAGC	1251
Oy	885	CATGCGCAGCAACTTTGATTTGTGTGAGATGACTTCAAGATGTGACAGCAAAATGA	944
Dp	1250	TGTTGAAAGAATCTAATTAATTTGTGTGAGACTTCAAGCGGTCAAGAAAGATGA	1191
Oy	945	ATCTGTGTTGTTGATTTTCTGTGGGGCTGTATCAAGCAACATCTATAGTTAGTTA	1004
Dp	1190	ATCGTTCTATGGAATTTTCTAGTTGAGCGGGAATTAAGCAACGTCAATGTTAGCTA	1131
Oy	1005	CAACCATCTGGAAAACAATGATGATATCTCTCGGCTCCACAACTTTCGGCTCAA	1064
Dp	1130	CAACCATTTGGGCAACAACAGATGAATGAATTTGCGCCCCCTCAAACTTCGGTCAA	1071
Oy	1065	GGAATCTCCAAAGCAACGTTGTTGACGATATGATCAACAGCAATGCCATCTCTATGA	1124
Dp	1070	GGAGATCTCCAAAGCAATGTATGTGACGACATGATGTGTAACAATTCATATATAGC	1011
Oy	1125	GCCTGTGAACATCCCGACCATGTTGTGTTAATTAAGTATGTGCTTAACGTAGGGATAG	1184
Dp	1010	GCCCGGTGAATCCGACACACGTTATTTGTATCAATGATGCTCCCTTATGTGGGTGACAG	951
Oy	1185	CAAGAGCCATGATGATCACTTCAAGATATTCATGGGTGAAAGAACACATGT	1244
Dp	950	CAAGAGGCAATGACAGATCACTTGGAGATCTTCATGGCGGAGATGAACAATCGT	891
Oy	1245	TTTGCAACAACATGTAGATTTCCCTTTTAACTGTCTCTATTTATCTTGAATTTGTCCT	1304
Dp	890	CATCAACAATACGTGCAAGATTTTAACTGTGCGCGCGCTCATCTTTGACATTTGATCTT	831

Oy	1305	TCCTTGCAGGCTGAGCACTAGAAATCCAGTTTAAAGCTGAAAATGAGGAAAATTCACATC	1364
Db	830	GCTGCGTAGTCTGCAACAAGAAATCCAATTCAGAAAGACGACAGAGAAAGTTCCATTC	771
Oy	1365	ATTCACACCGATTGCTACAAATTCAGACTAATCTACCAAGGCTCTCTGCTTCACCGGG	1424
Db	770	TTTTCACCCGCTTCCCGTACTCTCGAGCTACTACCAAGGCTCCCTTGGTACCCCAAG	711
Oy	1425	TACACCAAGTGTGAATSCAATTTGTCAAAGCAGCGTGCAATGCTGAAAAACATTAATGAGGC	1484
Db	710	TACCCCAAGTGTAAACCGTTGGCTAAGACAGAGGGCGATGCTAGAGATATCATGCGGC	651
Oy	1485	TTTGCTTGAATTTGGCCCAAGAAATACATGATTTCTCGAGTACAAAGT	1551
Db	650	GTGCAATTTGGGTGTTCCCCGACAAATTAACATGATGCTAGAAATACAAAGT	604

[illegible]

12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Query Match 45.6%; Score 698.8; DB 7; Length 988;
 Best Local Similarity 82.3%; Pred. No. 5.5e-200;
 Matches 802; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

177 TCCTCAAAATGGGGGATGATCTGTGGGTTGGGGTGAACAACAGGCTACCCACCG 236
 15 TTCCTCAAAATGGGGGATGATCTGTGGGTTGGGGTGAACAACAGGCTACCCAC 74
 237 TGGTGTATGCTTAAACAGAGAGACATTTTCATGGGCTACCAAGACAAGATTC 296
 75 AGGTGTATGCTTAAACAGAGAGATTTTCATGGGCTACCAAGATTCAGATC 134
 297 CAATTACTTGGGCTCCCTCACCACCGCTCAGCTATTCAGTTGATGATCTTCC 356
 135 CAATTACTTGGGCTCCCTCACCACCGCTCAGCTATTCAGTTGATGATCTTCC 194
 357 GGAATCTATGCTCCCTCACCACCGCTCAGCTATTCAGTTGATGATCTTCC 416
 135 GGAATCTATGCTCCCTCACCACCGCTCAGCTATTCAGTTGATGATCTTCC 254
 417 TGGGGGATGAGATATCAGCAACATGAACCTGCTGATGCTGATGCTGCTGCT 476
 255 TGGAGGATGAGATATCAGCAACATGAACCTGCTGATGCTGATGCTGCTGCT 314
 477 TGAATCTGATTTGCAAGACATTTGAGGCTTACATGATGATGATGATGATGAT 536
 315 TGAATCTGATTTGCAAGACATTTGAGGCTTACATGATGATGATGATGATGAT 374
 537 AATCTACGACCCGGAATTCATGCTGCTGCAACCAAGAGAGAGCTGCAACAG 596
 375 TATCTATGACCCGGAATTCATGCTGCTGCAACCAAGAGAGAGCTGCAACAG 434
 597 GGGCAACAAGAGAGAGAGATTCAGCAATTCATCAAGACATCAAGGCTTTA 656
 435 AGGAACCAAGAGAGAGAGATTCAGCAATTCATCAAGACATCAAGGCTTTA 494
 657 CACCAAGAGAGAGAGAGATTCAGCAATTCATCAAGACATCAAGGCTTTA 716
 495 CACCAAGAGAGAGAGAGATTCAGCAATTCATCAAGACATCAAGGCTTTA 554
 717 GGTGTAGGCTTAAATGACACCATGAGAGATCTCTGCTGCTGCTGCTGCTG 776
 555 GGTGTAGGCTTAAATGACACCATGAGAGATCTCTGCTGCTGCTGCTGCTG 614
 777 TGAATCTGCTTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 836
 615 TGAATCTGCTTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 674
 837 TGAATCTGCTTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 896
 675 CATGGAAGCCCTCAAGACATTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 734
 897 CACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 956
 735 CACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
 957 TGAATCTGCTTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1016
 795 TGAATCTGCTTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 854
 1017 AATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
 855 TGAATCTGCTTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 914
 1077 GAGCAAGCTTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1136
 915 AAGCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974

QY 1137 TCCGACCATGTTG 1150
 DB 975 CCTGACCATGTTG 988

RESULT 6

CK275352

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: APT TAG GTG ACA CTA TAG.

Location/Qualifiers

1..914
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POADM85"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DHI08-Tona"
 /clone_11b="potato abiotic stress cDNA library"
 /note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN

Query Match 42.5%; Score 651.4; DB 7; Length 914;
 Best Local Similarity 83.5%; Pred. No. 1.3e-185;
 Matches 739; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 649 AAGAAAGCACCACCAAGTGAAGAGTGTCTCTGTGACTGCAACACAGAGAGTAT 708
 DB 1 AAGAAAGCACCACCAAGTGAAGAGTGTCTCTGTGACTGCAACACAGAGATAC 60
 QY 709 AGCAATTTGTTGAGGCTTAATGACACCATGAGAAATCTTGGCTGCTGTGACAGA 768

Db	361	GCCTCAAAAACATTTGTTTCAGAGTCTTATGTGATTTGGCATTAAGAGAAACACTTAA	420
Qy	905	TTGGTGCAGATGACTTCAAGAGTGTGCAGACCAAAATGAAATCTGTGTGGTGAATTTTC	964
Db	421	TTGGTGGGATGACTTTAAGAGTGTGTCAAAACCAAGATGAAGTCGTGTGTGATTTTCC	480
Qy	965	TTGTGGGGGGCTGATCAAGCCAACTATAGTATAGTTAGTACCAACATCTGGGAAACAATG	1024
Db	481	TTGTTGGAGCTGGATTTAAGCCAAAGTCAAATAGTGAAGCTTACAAATCATTTGGGTAACTATG	540
Qy	1025	ATGGATGAATCTCTCGGCTCCACAAACCTTCCGCTCCAAAGAAATCTCCAAGACAAAG	1084
Db	541	ACGGATGAATCTTTCGTCTCCTCAAACTTCGGTCTTAAGAGATCTCAAAAAGTATAG	600
Qy	1085	TTGTGTGAGATATGCTCAACAGCAATGCGATCTCTATAGACTTGATGAACATCCGAC	1144
Db	601	TTGTGTGATGACATGGTTGCTAGCAACTCCATTTCTATGAAATCTGGCAGACCCCTGAC	660
Qy	1145	ATGTTGTGTATTATTAAGATATGCTTACGTAGGGGATAGCAAGAAGCCATGATGAGT	1204
Db	661	ATGTCGTGTATCAAGTATGTTCATATGTTGAGACAGAGAGGGCAATGATGAGT	720
Qy	1205	ACACTTGAGAGATATTCAATGGGTGGAAGAACACATGTTTGGACAACAATGTGAGG	1264
Db	721	ACATGTCAAGATTTTTCATGGGCGGAAAGACACATAGTTATGCAACAACCTGTGAGG	780
Qy	1265	ATTCCCTTTTAGCTGCTCCTATATCTTGAACCTTGTGCTTCTTGTGCTAGCTGAGACTA	1324
Db	781	ACTGCTCTTTTGGACGCTCCAAATTAATCTTGACCTTGTGCTTCTGCTGAACTCAGACATC	840
Qy	1325	GAATCCAGTTTAAAGCTGAAAATG--AGGAAAAATTCACATCTTCCACCACAGTTGCTAAC	1382
Db	841	GCATTTCAGACTCAAAAGCTGAAAAGGGAGGGAGAAAGTTCCACATCTTCCACCCTGTGTAC	900
Qy	1383	CATTCTCAGCTATCGACCAAGGCTCTCTGSGTTCC	1418
Db	901	TATTTCTCAGTACTACCAAGGCTCTCCCTGTGATCC	936

```

RESULT 8
CO085839
LOCUS
DEFINITION GR_C0085839 865 bp mRNA linear EST 16-JUN-2004
ACCESSION GR_Ea03E19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea03E19 5',
VERSION CO085839
KEYWORDS mRNA sequence.
SOURCE CO085839.1 GI:48776473
ORGANISM EST.
Gossypium raimondii
Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 865)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Wing,R.A., Kapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 03 row: E column: 19.
Location/Qualifiers
1..865

```

```
FEATURES
source
Location/Qualifiers
1..865
/organism="Gossypium raimondii"
/mol_type="rRNA"
/db_xref="taxon:29730"
/clone="GR_BA03E19"
/tissue_type="whole seedlings"
```

```

/dev_stage="first true leaves"
/lab_host="DH10b"
/clone_id="GR_Ea"
/note="Vector: PCW,SPORT-6.1; Site1: NotI; Site2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clones
plated/picked by AGI. More glycerol clones held in -80. "

```

Query Match	41.5%	Score 636.2;	DB 7;	Length 865;
Best Local Similarity	83.5%;	Pred. No. 5.2e-181;		
Matches 722;	Conservative 0;	Mismatches 143;	Indels 0;	Gaps 0

OY	426	GGAAATACAGCAACATGAACTGGCTGATGGCCAGGAGCAAGGATGTTGACATCGA	485
Db	1	GGACATTATGAGCATGAACTAGCTGATCCATAGGCTAGGAGCCAAAGTTTTCGACATGGA	60
OY	486	TTTGCAGAGCAGTTGAGGCGCTTACATGAAATCCATAGGTTCCACTCCCGGAAATCTACGA	545
Db	61	TTCTGCAAAAGCACTGAGACCCCTACATGGAATCCATAGGTTCCACTCCCTGGAATCTACGA	120
OY	546	CCCGGATTTTCATTGCTCTGCAACCAAGAGAGCGTGCACAACAAGTGTATTAAGGCACAA	605
Db	121	TTCTGATTTTCACTTGTCTGTACCAAGGTGAACGTGCCAATATGTATCAAGGGGACCA	180
OY	606	GCAAGACCAAGTTTCAGCAAAATCATCAAGAATCAGGCGTTTAAGAGCCACCAAGT	665
Db	181	GAAAGAACAAAGTTCCAGACAGTCTATCAAAAGCATCAAGAGATTCAAAGAGAAAACAAAGT	240
OY	666	GGAACAAGTGGTTGCTGCTGAGACCTGCAACACAGAGAGGATATGCAATTTGGTTTAGG	725
Db	241	GGACAAGGTTGTTTACTCTGACTGCAAAACACTGAGAGGTACAGCAATGTCATGCTGGG	300
OY	726	CCTTATGACACCATGAGAGATCTCTTGCGTCTGTGGAAGAAATGAGGCTGAGATTTTC	785
Db	301	GCTAAATGACACCGTGGAAAGCCTTATGCTTTTGGAAAGAAATGAATCAGAGATTTTC	360
OY	786	TCCTTTCCACTTGTATGCCATGCGCTGTGTGATGGAATAATCTTCCTTTCATTTATGGAAG	845
Db	361	TCCTTTCCACTTGTATGCTATGCTTGTGTGTTGTTGAAAAATGTTCTTTTCATCAAAATGCGAG	420
OY	846	CCCTCAGAACACTTTTGTACCAAGGCGTATGATCTTGCCATGCGAGAGAACACTTTGAT	905
Db	421	CCCACAAAACACTTTGTTCAGGGGTGATTTGATTTGGCTATTCAAAGAACGTGTGAT	480
OY	906	TGTGTGAGATGACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGGTTATTTTCT	965
Db	481	TGGAGAGATGACTTCAAGAGTGTGCGACAGCCAGATGAATCTGTCCCTCGTGAGATTTTCT	540
OY	966	TGTGGGGCGTGTATCAAGCCAACTATATGTTAGTTACAAACCATGCGGAAACAATGA	1025
Db	541	TGTTGGGGCGTGTATCAAGCCCAATGATATGATGATTCACACCATCTGGAAATATATGA	600
OY	1026	TGATATGATCTCTCGGCTCCACAACCTTCCGCTCCAGAGAAATCTCAAGAGCAACGT	1085
Db	601	TGGCATGATGTGTACAGACCACCAACCTTCCGTTCCAGAGAGATCTCAAGAGCAATGT	660
OY	1086	TGTTGAGATATGATGTCACACGAATGCACTCTCTATAGAGCTGTGTGAATATCCCGACA	1145
Db	661	TGTTATGATGACATGGTTTCAAGCAATGGAATCTGTATAGAGCTGTGTGAATATCTGTGATCA	720
OY	1146	TGTTGTTGTTATTAAGATATGTCCTTACGTAGGGGATAGCAAGAGACCAATGATGATGA	1205
Db	721	TGTTGTGTCTATCAAGTATGTGCCATATGTGGAGACAGCAAGAGACCATGATGATGTA	780
OY	1206	CACCTTCAGAGATATTCATGGGTGGAAGAACACCAATGTTTTCACAAACATGTGAGGA	1265
Db	781	CACATTCAGAGATATTCATGGGAGCAAGAACACCAATGTGTGTTGACAAACATGTGAGGA	840
OY	1266	TTCCCTTTTACTGCTCTCTATATTC	1290
Db	841	TTCCCTGTGTGCTGTCCCATTTATC	865

RESULT 9
LOCUS CK320178 906 bp mRNA linear EST 11-MAY-2004
DEFINITION L2P08a02 Populus stem seasonal library Populus deltoides cDNA, mRNA
sequence.
ACCESSION CK320178
VERSION CK320178.1 GI:47108601
KEYWORDS EST
SOURCE Populus deltoides
ORGANISM Populus deltoides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
1 (bases 1 to 906)
Park, S. and Han, K.-H.
Gene expression profile during seasonal growth cycle in poplar tree
Unpublished (2003)
JOURNAL
AUTHORS
TITLE
COMMENT
Contact: Kyung-Hwan Han
Department of Forestry
Michigan State University
126 Natural Resources, East Lansing, MI 48824-1222, USA
Tel: 517 353 4751
Fax: 517 432 1143
Email: hanky@msu.edu.
FEATURES
source
1..906
/organism="Populus deltoides"
/mol_type="mRNA"
/strain="IL-129"
/db_xref="taxon:3696"
/tissue_type="stem"
/dev_stage="1 year old"
/clone_lib="Populus stem seasonal library"
ORIGIN
Query Match 41.3%; Score 633; DB 7; Length 906;
Best Local Similarity 82.4%; Pred. No. 5e-180;
Matches 726; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
416 TTGGGGGATGGGATATTCAGCAACATGAACTGCTGATCCATGGCCAGGCAAAAGTGT 475
26 TTACGGCGCGGTCAATAGTGAATGAACTTGGCAGATGCCATGGCGGCAAGTTT 85
476 TTGACATCATGATTTGCAAGACGCTTGAGGCTTACATGGAATTCATGTTCCATCCCG 535
86 TCGACATTCATGCTGCAAAAGCAATGAGGCGCTACATGGAATTCATGCACTCCCTG 145
536 GAATCTAGACCGCGATTCATTTGCTGCGCAACAGAGAGGCTGCAACACGTGATTA 595
146 GAATTTACATCTCTGATTTTCATTTGCTGCGCAATCAAGGCTCAGTGTCAAAATGTTATCA 205
596 AAGGCAAAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAG 655
206 AAGGCACTTAAGAAAGCAAGTCACGCAAAATATGCAAAACATTAAGGAGTTCAAGAGA 265
656 CCACCAAGTGAACAAGGTGTTGCTCTGTGAATCCCAACAAGAGAGTATAGCAATT 715
266 AAAAAGAGTGAACAAGGTGTTGTTGTGATCTGCCAACAAGAGAGTACAGTAATA 325
716 TGGTGTAGGCGCTTAATGACCAATGAGAAATCTCTGCTGCTGTGACAGAAATGAGG 775
326 TTGTTGTGGGCTGAATGACCAATGAGAAACCTCTTGCTGCTGTGCGGAGAGAGTGAAT 385
776 CTGAGATTTCTCTTCCATCTGTATGACCATGCTGTGTGATGAGAAATGTTCCCTTTCA 835
386 CAGAGATATCCCATCACTCACTGTATGCTCTGGCTTGATTTATGAAATATTTCCCTTTCA 445
836 TTAATGAGAAAGCCCTCAAGAACATTTTGTACCAAGGCGTGAATGATCTTGCCATCGGAGGA 895
446 TAAATGAGAAAGCCCAAGAACATTTTGTTCAGGACTTGTGATTTGAGCTATTAAAGAGA 505
896 ACACCTTATTTGTGTGAGATGACTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTGG 955

Db
506 AAGTTTATTTGTGTGAGATGACTTTAAGATGTGTCAAGCCAAATGAATCTGTTCTGG 565
Qy 956 TTGATTTTCTTGTGGGGCGTGTATCAAGCCAAATCTATAGTTAGTTACAACATCTGG 1015
Db 566 TTGATTTCTTGTGGGGCGGTATTAAGCCAAACCTCAATAGTGAAGTTACAACCATCTGG 625
Qy 1016 GAAACAATGATGTATGAAATCTCTGGCTCCAAACCTTCGCTCCAAAGAAATCTCA 1075
Db 626 GAAACAATGATGTATGAAATCTATCAAGACCTTCAACCTTTGCTCAAGAAATCTCA 685
Qy 1076 AAGCAAGTGTGTGACATATGTCACAGCAATGCCATCTCTATGAGCTGTGTAAC 1135
Db 686 AAGCAATGTTTGTATGATCAATGCTCTCTAGCAATGCAATCTCTATGAGCTGTGTAAC 745
Qy 1136 ATCCGACCATGTTGTGTATTAAGTATGCTTACCTAGTGAAGGATAGCAAGAGCCA 1195
Db 746 ATCTGATCATATGTATGATTAATCAAGTATGTGCCATATGTTGAGATAGCAAGAGCTA 805
Qy 1196 TGGATGATGACCTTCAAGATATTCATATGCTGTGGAAGAAACCATGTTTGTGCAACA 1255
Db 806 TGGATGATGACCTTCAAGATATTCATATGCTGTGGAAGAAACCATGTTTGTGCAACA 865
Qy 1256 CATGTGAGATTCCTTTTACGCTGCTCTATTAATCTTGGAC 1296
Db 866 CCGTGTGAAGACTCCCTCTGCTGCTGACCAATCATCTCTGAC 906
RESULT 10
LOCUS COL13870 898 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_EB014121.r GR_EB Gossypium raimondii cDNA clone GR_EB014121
3', mRNA sequence.
ACCESSION COL13870
VERSION COL13870.1 GI:48812557
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 898)
Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
Wing, R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
JOURNAL
AUTHORS
TITLE
COMMENT
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 014 row: I column: 21.
FEATURES
source
1..898
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone_lib="GR_EB014121"
/tissue_type="Floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_libs="GR_EB"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clones
plated/picked by AGI. More glycerol clones held in -80."

Matches	729;	Conservative	0;	Mismatches	169;	Indels	0;	Gaps	0;
OY	145	GTCACTACCAATTTAAACCAACCCCAATGTTCCAAATTTGGGGTATGCTTGCGT	204						
Db	1	GTCAATATGAATTCAGACTGATACCCATGCCCTTAATTTGGGGGTATGCTTGCGG	60						
OY	205	TGGGGTGAACACGCGCTCTACCTCCACCGGTGTGTATTTGCTAACAGAGAGACATT	264						
Db	61	TGGGGAGGAAACATGTTTCAACCTCCACCGGTGTGTATTTAGCTAACAGAGGTATC	120						
OY	265	TGATGGGCTACAAAGACAGATTCACAAAGCCATTTACTTTGGCTCCTCAACCCAGCC	324						
Db	121	TCTTGGGCTACTAAGACAAAGGACAAAGGCTTAATTTCTTTGGTTCATTTGACTAAGA	180						
OY	325	TGAGCTATTCAGATTGATCTCTCCAGAGAGAGAAATCTATGCCCATTTCAAGAGCTG	384						
Db	181	TCAACGATCCGATTTGGGTCTTACATGAGAGAAATTTATGCTCATTAAAGACTCTT	240						
OY	385	CTTCCATATGTTAATCTGACGACATTTGTGTGGGGATGGGATATCAGCAACATGAC	444						
Db	241	CTTCCATATGTTAATCCCAATATGATTTGTGTGGAGATGGGACATTTAGTACATGAC	300						
OY	445	CTGGCTGATGCCATGGCCAGGCGCAAGGTGTTGACATGATTTGACAGAGAGATTGAG	504						
Db	301	CTAGCTGATGCAATGGCTAGGGCCAAAGGTTTTCGACATGATCTGCAAAAGCAACTG	360						
OY	505	CCTTACATGGAATCCATGTTCCATCCCGGAACTACGACCCGGATTTCAATGCTGCG	564						
Db	361	CCCTACATGGAATCCATGTTCCATCCCGGAACTACGATCTGATTTCAATGCTGCT	420						
OY	565	AAACCAAGAGAGCGTGCACCAACGATTTAAAGGCAACAGCAAGAGCAAGTTACGAA	624						
Db	421	AACCAAGAGAGAGCGTGCACCAATATGTCATCAAGAGGACCAAGAAAGCAAGTTAC	480						
OY	625	ATCATCAAGACATCAAGCGCTTTAAGAGAGCCCAAGTGTGAAGAGTGTGCTGCTG	684						
Db	481	GTATCAAAAGACATCAAGAGGATTTCAAGAGAAAGAGAGTGTGCTGCTGCTGCTG	540						
OY	685	TGAGCTGCCAAGACAGAGAGTATAGCAATTTGTTGAGCGCTTATGACACCATGAG	744						
Db	541	TGAGCTGCCAAGACAGAGAGTATAGCAATTTGTTGAGCGCTTATGACACCATGAG	600						
OY	745	AATCTCTGGCTGCTGTGACAGAAATGAGGCTGAGATTTCTCTTCCACTTTGATGCC	804						
Db	601	AACCTTATGCTTCTTTGAGAGAAATGATCAAGATTTCTCTTCCACTTTGATGCT	660						
OY	805	ATTGCTGTGTGATGAGAAATGTCCTTTCAATTAAGAGAGCCCTCAGAACTTTGTA	864						
Db	661	ATTGCTGTGTGATGAGAAATGTCCTTTCAATTAAGAGAGCCCTCAGAACTTTGTA	720						
OY	865	CCAGGGCTGATGATCTTCCATCGGAGAGAAACATTTGATTTGTTGAGATGACTTCAAG	924						
Db	721	CCAGGGCTGATGATTTGGCTATTCAAAGAACTGTCTATTTGGAGAGATGACTTCAAG	780						
OY	925	AGTGTGACAGCAAAATGAAATCTGTGTTGTTGATTTCTTGTGGGGCTGTATCAAG	984						
Db	781	AGTGTGACAGCAAAATGAAATCTGTGTTGTTGATTTCTTGTGGGGCTGTATCAAG	840						
OY	985	CCAACTCTATATGTTATGATCAACCATCTGGGAGAAACATGATGATGATCTTCGG	1042						
Db	841	CCAACTCTATATGTTATGATCAACCATCTGGGAGAAACATGATGATGATCTTCGG	898						

RESULT 11

BG044525

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

BG044525 670 bp mRNA linear EST 23-JUL-2004
 saae9e07.y1 Gm-c1059 Glycine soja cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-c1059-589 5' similar to TR:098SV4 098SV4
 MYO-INOSITOL-1-PHOSPHATE SYNTHASE.; mRNA sequence.
 BG044525
 BG044525.1 GI:12491342
 EST.
 SOURCE
 Glycine soja

ORGANISM Glycine soja
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
REFERENCE
 1 (bases 1 to 670)
 Shoemaker, R., Kaim, P., Vodkin, L., Espelid, J., Coryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Stepien, M., Thelning, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
TITLE
 Public Soybean EST Project
JOURNAL
 Unpublished (1999)
COMMENT
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 When it has been determined, an EST from the other end of this
 clone is listed in the 'Other ESTs on clone' field. This clone is
 available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
 High quality sequence stop: 422.
FEATURES
 Location/Qualifiers
 1..670
 /organism="Glycine soja"
 /mol_type="mRNA"
 /db_xref="taxon:3848"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1059-589"
 /issue_type="whole seedling, 2 week old, etiolated"
 /lab_host="DH10B"
 /clone_id="Gm-c1059"
 /note="Vector: plasmid pUC19. Site 1: EcoRI, Site 2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from 2 week old etiolated whole seedlings of P1468916.
 Complementary DNA was synthesized from mRNA using a primer
 consisting of a poly(dT) sequence with a XhoI restriction
 site. EcoRI adapters were ligated to the blunt-ended cDNA
 fragments followed by XhoI digestion. The cDNA fragments
 were directionally cloned into the EcoRI-XhoI restriction
 site of the plasmid vector. The ligated cDNA fragments
 were transformed into DH10B host cells (Gibco BRL). This
 library was constructed in the laboratory of Dr. Randy
 Shoemaker at Iowa state university."

ORIGIN

Query Match 40.8%; Score 625.8; DB 4; Length 670;

Best Local Similarity 96.9%; Pred. No. 6.8e-178; Mismatches 18; Indels 3; Gaps 1;

Matches	649;	Conservative	0;	Mismatches	18;	Indels	3;	Gaps	1;
OY	474	GTTTACATCGATTTTGCAAGACAGTGTAGGCTTACATGCAATGATGTTCCCTCC	533						
Db	1	GTTTACATCGATTTTGCAAGACAGTGTAGGCTTACATGCAATGATGTTCCCTCC	60						
OY	534	CGGATCTACGACCGGATTTTCAATGCTGCCAACCAAGAGAGCGTCCCAACAGTAT	593						
Db	61	CGGATCTACGACCGGATTTTCAATGCTGCCAACCAAGAGAGCGTCCCAACAGTAT	120						
OY	594	TAAAGGCACAAAGCAAGACAGTTCAGCAATCATCAAGACATCAAGCGTTTAAAGA	653						
Db	121	TAAAGGCACAAAGCAAGACAGTTCAGCAATCATCAAGACATCAAGCGTTTAAAGA	180						
OY	654	AGCCACCAAGTGAAGCAAGGTGTTGCTGTGAGCTCCCAACAGAGAGATATGACA	713						
Db	181	AGCCACCAAGTGAAGCAAGGTGTTGCTGTGAGCTCCCAACAGAGAGATATGACA	240						
OY	714	TTTGGTGTAGGCTTAAATGACACCATGAGAGATCTTGGCTGTGAGCAGAAATGA	773						
Db	241	TTTGGTGTAGGCTTAAATGACACCATGAGAGATCTTGGCTGTGAGCAGAAATGA	300						

ORGANISM *Vitis vinifera*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 842)
 Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Leslie, A., Xu, J.,
 Jones, K. and Cook, D.
 Expressed sequence tags from *Vitis vinifera* 'Cabernet sauvignon'
 berries at various developmental stages
 Unpublished (2003)

TITLE
 JOURNAL
 COMMENT
 CONTACT: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: ACGGTACCGACATATGCC.
 Location/Qualifiers
 1..842
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultiivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAB20007_I1a_Fa_D05"
 /sex="Hermaphrodite"
 /dev_stage="Bloom"
 /clone_lib="Cabernet Sauvignon Flower bloom - CAB2"
 /note="Organ: Flower - Bloom; Vector: pDNR; Site_1: S11;
 Site_2: S11; CAB2 is a cDNA library of *Vitis vinifera* cv.
 'Cabernet Sauvignon' Clone 8 berries. Samples were
 collected at full bloom (80 to 100 flowers showing
 dehiscence of calypters or caps and anthers fully
 extended). Sampled vines were located at the University of
 California, Davis, Experimental Vineyard. cDNAs were made
 by oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGAGATGCGCATTAAGCGCGG-3' and
 5'-AAGTGAAGCGGAGCGGCGGAGCAATG-3' (30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 40.7%; Score 623.6; DB 7; Length 842;
 Best Local Similarity 84.6%; Pred. No. 3.5e-177;
 Matches 712; Conservative 0; Mismatches 129; Indels 1; Gaps 1;

382 CTGCTTCCATGCTTAATCTGACGACATTGTGT-TTGGGGGATGGATATGACCAACT 440
 1 CTCCTCCCTATGTAACCCAGANTGACATTGGATCTTGGGGATGGACATCATGACAT 60
 441 GAACCTGGGTATGTCATGGCGGCGGCAAGGTTTTCATCATGATTTGCGAAGCAGTT 500
 61 GAACCTGGAGATGTAATGGCCAGGCTTAAGGTCCTGACATGATTTGCGAAGCAGTT 120
 501 GAGGCTTATCATGATGATCATGCTTCCATCTCCCGGAACTCTACGACCCGGAATTTGATTC 560
 121 GAGGCTTATCATGATGATCATGCTTCCATCTCCCGGAACTCTACGACCCGGAATTTGATTC 180
 561 TGCCCAACCAAGAGAGCGCTGCGCAACAGCTGATTAAAGGCGACAAAGCAAGCAAGATTCA 620
 181 TGCCCAACCAATCTCTCGGCGCAACAGCTGATTCAAGGCGACCAAGAAAGAGAGCTGCA 240
 621 GCAATTCATCAAGACATCAAGCGGTTTAAAGAACCCCAAGGCGCAAGGCTGTTGT 680
 241 GAGATCATTAAGATTAATTAAGGAGTTTAAGGAGCAACCAAGGTGAGTAATGATGAT 300
 681 CCGTGAGCTGCGCAACAGAGGATATGACAAATTTGTTAGCGCTTAATGACACCAT 740
 301 GGTGTGAGCTGCGCAACAGAGGATATGACAAATTTGTTAGCGCTTAATGACACCAT 360
 741 GGAGATCTCTTGGCTGCTGTGAGCAAGAAATGAGGCTGAGATTTCTCTTCCACCTTGT 800
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 361 AGAAATCTCTTGGCTTCTCTGATAGGAATGAGCTGAGATCTTCTCCCTCCACTTGTGA 420
 Qy 801 TGCCATTCCTGCTGTGATGGAATAATGTTCCCTTTCATTAATGGAAGCCCTCAGAACACTTT 860
 Db 421 TGCTCTGCTTGTGTATGAAAATGTTCCCTTATCAATGGAAGCCCAAGAACACTTT 480
 Qy 861 TGTAACAGGGCTGATGATGATCTTGGCCATCGCAGGAACAATTGATGTTGGTGAATGACTT 920
 Db 481 TGTTCAGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 Qy 921 CAAGAGTGTGACAGCAAAATGGAATCTGCTGTTGATTTTCTGTTGGGGCTGGAT 980
 Db 541 TTAAGATGTCTAAACCAAGATGAATCTGATGATGATGATGATGATGATGATGATGATGAT 600
 Qy 981 CAAGCAACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1040
 Db 601 CAAGCAACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 Qy 1041 GGCTTCACAAACCTTCGGCTTCCAAAGAAATCTCCAAAGCAACCTTGTGACATATGCT 1100
 Db 661 TGCTCTCAAAACCTTCGGCTTCCAAAGAAATCTCCAAAGCAACCTTGTGACATATGCT 720
 Qy 1101 CAACAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1160
 Db 721 TTCCAGCAACGCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 Qy 1161 GTATGCTCTTACGTAAGGAGATAGCAAGAGCCATGATGATGATGATGATGATGATGATGAT 1220
 Db 781 GTATGCTCTTACGTAAGGAGATAGCAAGAGCCATGATGATGATGATGATGATGATGATGAT 840
 Qy 1221 CA 1222
 Db 841 CA 842

RESULT 14
 CB973407
 LOCUS
 DEFINITION
 Vitis vinifera cDNA clone CAB30002_Ic_Fc_D10 5', mRNA sequence.
 CB973407
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 867)
 Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Leslie, A., Xu, J.,
 Jones, K. and Cook, D.
 Expressed sequence tags from *Vitis vinifera* 'Cabernet sauvignon'
 berries at various developmental stages
 Unpublished (2003)
 CONTACT: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: ACGGTACCGACATATGCC.
 Location/Qualifiers
 1..867
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultiivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAB30002_Ic_Fc_D10"
 /sex="Hermaphrodite"
 /dev_stage="Berry stage I"
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Berry Stage I - CAB3"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 16:00:16 ; Search time 843.546 Seconds
(without alignments)
10758.117 Million cell updates/sec

Title: US-10-718-952-11
Perfect score: 1533
Sequence: 1 agcttcacgcagaattttaa.....tgattctcgagtaacaagcga 1533

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1533	100.0	1533	12	ADQ14500	Adq14500 Mutant so
2	1533	100.0	1533	13	AD882001	Ad882001 Soybean m
3	1531.4	99.9	1533	12	ADQ14504	Adq14504 Wild type
4	1531.4	99.9	1533	12	ADQ14502	Adq14502 Mutant so
5	1531.4	99.9	1533	13	AD882005	Ad882005 Soybean m
6	1531.4	99.9	1533	13	AD882003	Ad882003 Soybean m
7	1464.2	95.5	1533	12	ADQ14498	Adq14498 Mutant so
8	1464.2	95.5	1533	13	AD881999	Ad881999 Soybean m
9	1464.2	95.5	1533	13	AD881993	Ad881993 Soybean m
10	1464.2	95.5	1760	12	ADQ14490	Adq14490 Wild type
11	1464.2	95.5	1782	2	AAV62440	AAV62440 Soybean w
12	1462.6	95.4	1533	2	AAV62443	AAV62443 Soybean m
13	1462.6	95.4	1533	2	ADQ14494	Adq14494 Mutant so
14	1462.6	95.4	1533	13	AD881997	Ad881997 Soybean m
15	1097.8	71.6	1950	2	AAx30402	AAx30402 Nicotiana
16	1000.2	65.2	1781	4	AAc87643	AAc87643 Brassica
17	998.6	65.1	1533	6	ABz13633	ABz13633 Arabidops
18	998.6	65.1	1533	12	ADN73524	ADN73524 Thale cre
19	998.6	65.1	1938	3	AAc50242	AAc50242 Arabidops
20	987.2	64.4	1536	3	AAc34428	AAc34428 Arabidops

21	985.8	64.3	1665	2	AAx09006	AAx09006 Inducible
22	985.6	64.3	1837	3	AAc34806	AAc34806 Arabidops
23	984.2	64.2	1959	4	AAc85922	AAc85922 MIP synth
24	979.4	63.9	1931	2	AAx24407	AAx24407 Maize mdo
25	941.6	61.4	1759	6	ABQ72653	ABQ72653 Human MBD
26	821.4	53.6	1536	12	ADP43918	ADP43918 P. coarct
27	782	51.0	1719	3	AAc51095	AAc51095 Arabidops
28	547.2	35.7	1605	8	ABT19345	ABT19345 Aspergill
29	547.2	35.7	1605	8	ABT21165	ABT21165 Aspergill
30	527.2	34.4	2280	4	ABU17361	ABU17361 Drosophila
31	499	32.6	1772	4	AAH14171	AAH14171 Human CDN
32	499	32.6	1818	4	AAH15251	AAH15251 Human CDN
33	499	32.6	1825	6	ABT07176	ABT07176 Human ova
34	499	32.6	1852	4	AAH47741	AAH47741 Human hmi
35	499	32.6	1862	4	AAI59232	AAI59232 Human pol
36	499	32.6	2380	13	ADQ86789	ADQ86789 Human tum
37	499	32.6	2380	13	ADQ85720	ADQ85720 Tumour-as
38	499	32.6	2380	13	ACN40520	ACN40520 Human tum
39	486.6	31.7	1991	12	ADG73735	ADG73735 Aspergill
40	478.2	31.2	1833	4	AAH47740	AAH47740 Murine mm
41	464.4	30.3	618	13	ACN62990	ACN62990 Cotton ca
42	453.6	29.6	1835	12	ADQ25429	ADQ25429 Human sof
43	451.2	29.4	1702	10	ADB62765	ADB62765 Human CDN
44	450.8	29.4	2338	4	AAI61018	AAI61018 Human pol
45	434.2	28.3	1704	8	ABT20567	ABT20567 Aspergill

ALIGNMENTS

RESULT 1	ADQ14500	standard; cDNA; 1533 BP.
ID	ADQ14500;	(first entry)
AC	23-SEP-2004	
XX		
DE	Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.	
XX		
KW	soybean; myo-inositol 1-phosphate synthase; gene; ss;	
KW	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;	
KW	inorganic phosphate; mutant.	
XX		
OS	Glycine max.	
OS	Synthetic.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..1533
FT		/*tag= a
FT		/product= "Mutant soybean myo-inositol 1-phosphate synthase #3"
XX		
PN	US2004128713-A1.	
XX		
PD	01-JUL-2004.	
XX		
PF	21-NOV-2003; 2003US-00718952.	
XX		
PR	08-APR-1997; 97US-00835751.	
PR	07-APR-1998; 98MO-US006822.	
PR	26-APR-1999; 99US-00293315.	
XX		
PR	11-MAR-2002; 2002US-00025003.	
XX		
PA	(HITZ/) HITZ W D.	
PA	(SEBA/) SEBASTIAN S A.	
PA	(GRAC/) GRACE D J.	
PA	(STRE/) STREIT L G.	
XX		
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
XX		
DR	WPI: 2004-533135/51.	
DR	P-PSDB; ADQ14501.	

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Claim 10; SEQ ID NO 11; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 100.0%; Score 1533; DB 12; Length 1533;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGTTCATGAGAAATTTTAAAGTAGAGATCCTTAATGTGAATGACCCGAGCTGAGATT 60
DB 1 AAGTTCATGAGAAATTTTAAAGTAGAGATCCTTAATGTGAATGACCCGAGCTGAGATT 60
QY 61 CAGTCCGCTGACAACTACGAAACACCGAACTTGTTCAGAGAACAGAAATGGCACTTAT 120
DB 61 CAGTCCGCTGACAACTACGAAACACCGAACTTGTTCAGAGAACAGAAATGGCACTTAT 120
QY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCATCACTAATTTAAACCAACCCATGTTCCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCATCACTAATTTAAACCAACCCATGTTCCA 180
QY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCATCACTAATTTAAACCAACCCATGTTCCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCATCACTAATTTAAACCAACCCATGTTCCA 180
QY 181 AATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCCCTGACCGGTGAT 240
DB 181 AATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCCCTGACCGGTGAT 240
QY 181 AATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCCCTGACCGGTGAT 240
DB 181 AATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCCCTGACCGGTGAT 240
QY 241 GTTATTTGCTAACGAGAGACATTTTCATGGGCTACAAAGAACAGATTCACAAAGCCAT 300
DB 241 GTTATTTGCTAACGAGAGACATTTTCATGGGCTACAAAGAACAGATTCACAAAGCCAT 300
QY 301 TACTTTGGCTCCCTCAACCCCAAGGCTCAGCTAATTCAGTTGATCTTCCAGGAGAGGAA 360
DB 301 TACTTTGGCTCCCTCAACCCCAAGGCTCAGCTAATTCAGTTGATCTTCCAGGAGAGGAA 360
QY 361 ATCTATGCCCCATTCAAGAGTCTGCTTCCAAATGTTTAACTCTGACACATTTGTTTGGG 420
DB 361 ATCTATGCCCCATTCAAGAGTCTGCTTCCAAATGTTTAACTCTGACACATTTGTTTGGG 420
QY 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCAATGACGAGGCAAGGATTTTGAAC 480
DB 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCAATGACGAGGCAAGGATTTTGAAC 480
QY 481 ATCGATTTTCAGAAAGCAGTTGAGCCCTTACATGGAATCCATGCTTCCCTCCGGAATC 540
DB 481 ATCGATTTTCAGAAAGCAGTTGAGCCCTTACATGGAATCCATGCTTCCCTCCGGAATC 540
QY 541 TACGACCCGAGATTTCATTGCTGCAACCAAGAGAGCGTGCACCAACGATGATTAAGGCG 600
DB 541 TACGACCCGAGATTTCATTGCTGCAACCAAGAGAGCGTGCACCAACGATGATTAAGGCG 600

QY 601 ACAAAGCAGAGCAAGTTCAAGCAATATCAAGACATCAAGGCGTTTAAAGAACCAACC 660
DB 601 ACAAAGCAGAGCAAGTTCAAGCAATATCAAGACATCAAGGCGTTTAAAGAACCAACC 660
QY 661 AAAGTGACAAAGGTGTTGCTCTGTGACCTGCCAACACAGAGAGTATAGCAATTTGGTT 720
DB 661 AAAGTGACAAAGGTGTTGCTCTGTGACCTGCCAACACAGAGAGTATAGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTCCACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 ATTTCTCTCCACCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTTGCCATGCGAGAACACT 900
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTTGCCATGCGAGAACACT 900
QY 901 TTGATTTGGTGAGATGACCTTCAGAGTGTGACAGCAACAAATGAAATCTGTGTTGAT 960
DB 901 TTGATTTGGTGAGATGACCTTCAGAGTGTGACAGCAACAAATGAAATCTGTGTTGAT 960
QY 961 TTTCTTGTGGGGGCTGTGATCAAGCCAACTATATAGTTAGTTTCAACCATCTGGGAAAC 1020
DB 961 TTTCTTGTGGGGGCTGTGATCAAGCCAACTATATAGTTAGTTTCAACCATCTGGGAAAC 1020
QY 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 AAGCTTTGTGACGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 AAGCTTTGTGACGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GACCAATGTTGTTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GACCAATGTTGTTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTTATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 GAGTACACTTCAGAGATATTTATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 GAGATTTCCCTTTTGAAGTCTGCTCTTATTTCTTGAATGATGATGATGATGATGAT 1320
DB 1261 GAGATTTCCCTTTTGAAGTCTGCTCTTATTTCTTGAATGATGATGATGATGATGAT 1320
QY 1321 ACTAGAAATCAAGTTTAAAGCTGAAATGAGGAAATTCACATCTCATTCACCCAGTTGCT 1380
DB 1321 ACTAGAAATCAAGTTTAAAGCTGAAATGAGGAAATTCACATCTCATTCACCCAGTTGCT 1380
QY 1381 ACCATTTCTCAGTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACAGTGTGAT 1440
DB 1381 ACCATTTCTCAGTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACAGTGTGAT 1440
QY 1441 GCATTGTCAAAAGCAGCTGCAATGCTGGAATAACATATGAGGAGCTTGTGTTGATG 1500
DB 1441 GCATTGTCAAAAGCAGCTGCAATGCTGGAATAACATATGAGGAGCTTGTGTTGATG 1500
QY 1501 CCAGAGAAATTAACATGATTTCTGAGTACAGTGA 1533
DB 1501 CCAGAGAAATTAACATGATTTCTGAGTACAGTGA 1533

RESULT 2
ADS82001
ID ADS82001 standard; cDNA; 1533 BP.
XX
XX ADS82001;
AC
XX
XX
DT 18-NOV-2004 (first entry)
XX

DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
XX
XX Soybean, plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
OS Glycine max; line 29010CP01.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
FT mutation replace(260,G)
FT /tag= b
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
XX P-PSDB; ADS82002.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 11; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homologous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a mutant myo-
CC inositol 1-phosphate synthase.
XX
XX
SQ Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 100.0%; Score 1533; DB 13; Length 1533;
Beet Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTCATCGAATAATTTTAAGGTAGAGAGTCTTAATGTGAATACACCGAGCTGAGATT 60
|||
DB 1 ATGTCATCGAATAATTTTAAGGTAGAGAGTCTTAATGTGAATACACCGAGCTGAGATT 60
QY 61 CAGTCGCGTGTACAACTACGAAACGACCGAATCTGTTCACGAGAACGAAATGGACCTAT 120
|||
DB 61 CAGTCGCGTGTACAACTACGAAACGACCGAATCTGTTCACGAGAACGAAATGGACCTAT 120
QY 121 CAGTGGATTGTCAAAACCCAAATCCGTCACTACCAATTTAAACCAACACCCATGTCCA 180
|||
DB 121 CAGTGGATTGTCAAAACCCAAATCCGTCACTACCAATTTAAACCAACACCCATGTCCA 180
QY 181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGAACACAGGCTCTACCTCACCCTGTGT 240
|||
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGAACACAGGCTCTACCTCACCCTGTGT 240
QY 241 GTTATTTCTPACAGAGAGAGACATTTCAATGGGCTACAAAGAGAGACAGATTCAACAGCCAT 300
|||
DB 241 GTTATTTCTPACAGAGAGAGACATTTCAATGGGCTACAAAGAGAGACAGATTCAACAGCCAT 300
QY 301 TACTTGGGCTCCCTCAGCCAAAGCCTCAGCTATTTCGAGTTGATTCCTTCCAGGAGAGAA 360
|||
DB 301 TACTTGGGCTCCCTCAGCCAAAGCCTCAGCTATTTCGAGTTGATTCCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCATGTGTTAATCCTGACGACATTTGTGGG 420
|||
DB 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCATGTGTTAATCCTGACGACATTTGTGGG 420
QY 421 GGATGGGATATAGCAAGATGAACCTGGCTGATGTGCATGGCCAGGAGCAAGGTTTGAAC 480
|||
DB 421 GGATGGGATATAGCAAGATGAACCTGGCTGATGTGCATGGCCAGGAGCAAGGTTTGAAC 480
QY 481 ATCGATTTGCAAGAACAGTTGAGGCTTTACATGATCATGATGTTCCACTCCCGGAATC 540
|||
DB 481 ATCGATTTGCAAGAACAGTTGAGGCTTTACATGATCATGATGTTCCACTCCCGGAATC 540
QY 541 TAGACACCGGATTTTATTTGCTGCTCCACCAAGAGAGCGTCCAAACGTTGTAAGGCG 600
|||
DB 541 TAGACACCGGATTTTATTTGCTGCTCCACCAAGAGAGCGTCCAAACGTTGTAAGGCG 600
QY 601 ACAAAGCAAGAGCAAGTTTCAAGAAATCATCAAGATCAAGAGCGTTTAAAGAAACCAAC 660
|||
DB 601 ACAAAGCAAGAGCAAGTTTCAAGAAATCATCAAGATCAAGAGCGTTTAAAGAAACCAAC 660
QY 661 AAAGTGACAAAGGTGGTGTCTGTGACCTGCAACACAGAGAGGTATAGCAATTTGGTT 720
|||
DB 661 AAAGTGACAAAGGTGGTGTCTGTGACCTGCAACACAGAGAGGTATAGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACGATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAG 780
|||
DB 721 GTAGGCTTTAATGACACGATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACTTGTATGCAATTCCTGTGTGATGAGAAATGTTCTTCAATAT 840
|||
DB 781 ATTTCTCTTCCACTTGTATGCAATTCCTGTGTGATGAGAAATGTTCTTCAATAT 840
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGTGATCTTGCATGCGGAGAACCT 900
|||
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGTGATCTTGCATGCGGAGAACCT 900
QY 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTGTTGAT 960
|||
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTGTTGAT 960
QY 961 TTTCTTGTGGGGGCTGTGTATCAAGCAACATCTATAGTTAGTTACACATCTGTGGAAC 1020
|||
DB 961 TTTCTTGTGGGGGCTGTGTATCAAGCAACATCTATAGTTAGTTACACATCTGTGGAAC 1020
QY 1021 AATGATGATGATATCTCTCGGCTCCAAACCTTCCGCTCCAGAGAAATCTCCAAAGAC 1080
|||
DB 1021 AATGATGATGATATCTCTCGGCTCCAAACCTTCCGCTCCAGAGAAATCTCCAAAGAC 1080

```
QY 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCC 1140
DB 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCC 1140
QY 1141 GACCATGTTGTTGTTATTTAAGTATGCTCTTACGTAAGGAGATAGCAAGAGACCATGAT 1200
DB 1141 GACCATGTTGTTGTTATTTAAGTATGCTCTTACGTAAGGAGATAGCAAGAGACCATGAT 1200
QY 1201 GAGTACACTTCAAGATATTTCAATGGGTGGAAGAACACCATTTGTTGCAACAACATGT 1260
DB 1201 GAGTACACTTCAAGATATTTCAATGGGTGGAAGAACACCATTTGTTGCAACAACATGT 1260
QY 1261 GAGGATTTCCCTTTTACGCTCTCTATTTATCTTGACATTTGCTCTTCTTGACGCTGAC 1320
DB 1261 GAGGATTTCCCTTTTACGCTCTCTATTTATCTTGACATTTGCTCTTCTTGACGCTGAC 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAAATGAGGGAATTCACATTCATTCACCCAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAAATGAGGGAATTCACATTCATTCACCCAGTTGCT 1380
QY 1381 ACCATTCTGAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACAGGTGAT 1440
DB 1381 ACCATTCTGAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACAGGTGAT 1440
QY 1441 GCATTGTCAAGGAGCGTCAATGCTGGAATAATATGAGGCTTTGTTGATTTGGCC 1500
DB 1441 GCATTGTCAAGGAGCGTCAATGCTGGAATAATATGAGGCTTTGTTGATTTGGCC 1500
QY 1501 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1533
DB 1501 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1533
```

```
RESULT 3
ADQ14504 standard; cDNA; 1533 BP.
```

```
ADQ14504;
```

```
23-SEP-2004 (first entry)
```

```
Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.
```

```
Soybean; myo-inositol 1-phosphate synthase; gene; ss;
myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
inorganic phosphate.
```

```
Glycine max.
```

```
Location/Qualifiers
```

```
Key 1.1533
FT CDS /tag=a
FT /product="wild type soybean myo-inositol 1-phosphate
FT synthase #2"
```

```
US2004128713-A1.
```

```
01-JUL-2004.
```

```
21-NOV-2003; 2003US-00718952.
```

```
08-APR-1997; 97US-00835751.
```

```
07-APR-1998; 98WO-US006822.
```

```
26-APR-1999; 99US-00299315.
```

```
11-MAR-2002; 2002US-00025003.
```

```
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
```

```
XX WPI; 2004-533135/51.
DR P-PsDB; ADQ14505.
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX Claim 4; SEQ ID NO 15; 48bp; English.
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
```

```
Query Match 99.9%; Score 1531.4; DB 12; Length 1533;
```

```
Best Local Similarity 99.9%; Pred. No. 0;
```

```
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGTCATCGAGATTTTAAAGTAGAGTCTTAATGTAAGTACCGAGATCGAGATT 60
DB 1 ATGTCATCGAGATTTTAAAGTAGAGTCTTAATGTAAGTACCGAGATCGAGATT 60
QY 61 CAGTCGGTGTCAACTAGAGAAACCGAAGCTGTGTCAGGAACAGAAATGGACCTAT 120
DB 61 CAGTCGGTGTCAACTAGAGAAACCGAAGCTGTGTCAGGAACAGAAATGGACCTAT 120
QY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCAACTACCAATTTAAACCAACCATTTTCCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCAACTACCAATTTAAACCAACCATTTTCCA 180
QY 181 AAATTGGGGGTGATGCTTGTGGGTGGGAAACAACGGCTCTACCCCTACCGGTGT 240
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGAAACAACGGCTCTACCCCTACCGGTGT 240
QY 241 GTTATTTGTCAACAGAGAGACATTTCAATGGGCTCAAGAGCAAGATTCACCAAGCAAT 300
DB 241 GTTATTTGTCAACAGAGAGACATTTCAATGGGCTCAAGAGCAAGATTCACCAAGCAAT 300
QY 301 TACTTTGGCTCCCTCAACCCCAAGCTCAGCTATTGAGATTGATCTTCCAGGAGAGANA 360
DB 301 TACTTTGGCTCCCTCAACCCCAAGCTCAGCTATTGAGATTGATCTTCCAGGAGAGANA 360
QY 361 ATCTATGCCCATTCAGAGAGTCTGCTCAATGTTAACTCTGACGACATTTGTGGG 420
DB 361 ATCTATGCCCATTCAGAGAGTCTGCTCAATGTTAACTCTGACGACATTTGTGGG 420
QY 421 GATGAGATTTCAAGAAATGAACTGGCTGATGCAATGGGCAAGGAAAGTGTTCAC 480
DB 421 GATGAGATTTCAAGAAATGAACTGGCTGATGCAATGGGCAAGGAAAGTGTTCAC 480
QY 481 ATCGATTTGCAAGAAAGTTGAGGCTTTACATGGAATTCATGATTTCCCTCCCGGAATC 540
DB 481 ATCGATTTGCAAGAAAGTTGAGGCTTTACATGGAATTCATGATTTCCCTCCCGGAATC 540
QY 541 TACGACCCGATTTTCATTTCTGCCAACCAAGAGAGCGTGCACCAACGATTAAGGGC 600
```

Db 541 TACGACCGGATTTCTTGTGCTGCAACAGAGAGAGCGTCCAAACACGATTAAGGGC 600
Qy 601 ACAAGCAAGAGCAAGTTGACGAATCATCAAGACATCAAGGCGTTTAAAGAGCCACC 660
Db 601 ACAAGCAAGAGCAAGTTGACGAATCATCAAGACATCAAGGCGTTTAAAGAGCCACC 660
Qy 661 AAGTGAAGAGGAGTGTGCTGCTGTGAGACGCAACACAGAGAGGATATGCAATTTGGTT 720
Db 661 AAGTGAAGAGGAGTGTGCTGCTGTGAGACGCAACACAGAGAGGATATGCAATTTGGTT 720
Qy 721 GTAGGCTTAATGACACCAATGAGAAATCTTGGCTGTGTGACAGAAATGAGGCTGAG 780
Db 721 GTAGGCTTAATGACACCAATGAGAAATCTTGGCTGTGTGACAGAAATGAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCAATGCGCTGTGTGATGGAATAATGTTCTTTCAATTA 840
Db 781 ATTTCTCTTCCACCTTGTATGCAATGCGCTGTGTGATGGAATAATGTTCTTTCAATTA 840
Qy 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATGATTTGCCATCGGAGAACACT 900
Db 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATGATTTGCCATCGGAGAACACT 900
Qy 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAACCAATCTGTGTTGTTGAT 960
Db 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAACCAATCTGTGTTGTTGAT 960
Qy 961 TTTCTTGTGGGGCTGTGTATCAAGCAACATCTATAGTTAGTTACCAACCATCTGGAAAC 1020
Db 961 TTTCTTGTGGGGCTGTGTATCAAGCAACATCTATAGTTAGTTACCAACCATCTGGAAAC 1020
Qy 1021 AATGATGATGATGATCTTCTGCTGCTCAACAACTTCCGCTCCAGAGAAATCTTCAAGAGC 1080
Db 1021 AATGATGATGATGATCTTCTGCTGCTCAACAACTTCCGCTCCAGAGAAATCTTCAAGAGC 1080
Qy 1081 AACGTTGTGACGATATGATGCAACAGCAATGCAATCTGTATAGGCTGTGTAACATCCC 1140
Db 1081 AACGTTGTGACGATATGATGCAACAGCAATGCAATCTGTATAGGCTGTGTAACATCCC 1140
Qy 1141 GACCATGTTGTTGTTATTAAGATAGTCCCTTACGTAAGGAGATGACAGAGCCATGAT 1200
Db 1141 GACCATGTTGTTGTTATTAAGATAGTCCCTTACGTAAGGAGATGACAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAAGATATTTCAATGAGGAGAAACCACTTTTGTGACACACATATG 1260
Db 1201 GAGTACACTTCAAGATATTTCAATGAGGAGAAACCACTTTTGTGACACACATATG 1260
Qy 1261 GAGGATTTCCCTTTAGCTGCTCTATTTATCTTGAACCTTGTGCTGAGCTGAGC 1320
Db 1261 GAGGATTTCCCTTTAGCTGCTCTATTTATCTTGAACCTTGTGCTGAGCTGAGC 1320
Qy 1321 ACTAGATCAAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACACCAAGTTGCT 1380
Db 1321 ACTAGATCAAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACACCAAGTTGCT 1380
Qy 1381 ACCATTTCTAGCTATTTGACCAAGGCTCTCTGTGTTCCACCGGTTACACCAAGTTGAT 1440
Db 1381 ACCATTTCTAGCTATTTGACCAAGGCTCTCTGTGTTCCACCGGTTACACCAAGTTGAT 1440
Qy 1441 GCATTTGCAAAAGAGGCTGCAATGCTGAAAACATTAATGAGGCTGTGTGATTTGGCC 1500
Db 1441 GCATTTGCAAAAGAGGCTGCAATGCTGAAAACATTAATGAGGCTGTGTGATTTGGCC 1500
Qy 1501 CCAGAGATATACATGATTTCTGAGTACCAAGTGA 1533
Db 1501 CCAGAGATATACATGATTTCTCGAGTACCAAGTGA 1533

RESULT 4
ADQ14502
ID ADQ14502 standard; cDNA; 1533 BP.
XX
AC ADQ14502;

XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
XX
KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
OS Glycine max.
OS Synthetic.
FH Key location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #4"
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
DR WPI; 2004-53335/51.
XX
DR P-PsDB; ADQ14503.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Example 8; SEQ ID NO 13; 48bp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 99.9%; Score 1531.4; DB 12; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGTCATCGAGATTTTAAGGTAGAGTCTCTATATGAGTACACGAGACGTGAGATT 60
|||||


```

Db      1 ATGTCATCGAAGATTTTAAAGTAGAGTCCCTAATGTGAATGACACCGAGACTGAGAT 60
Qy      61 CAGTCGGTGTACAACTACGAAACCAACCGAATCTGTTCACGAGACAGGAATGGACCTAT 120
Db      61 CAGTCGGTGTACAACTACGAAACCAACCGAATCTGTTCACGAGACAGGAATGGACCTAT 120
Qy      121 CAGTGGATTTGTCAAAACCCAAATCCGTCAACTACCAATTTAAACCAACCCATGTTTCA 180
Db      121 CAGTGGATTTGTCAAAACCCAAATCCGTCAACTACCAATTTAAACCAACCCATGTTTCA 180
Qy      181 AAATTTGGGGGTATGCTTGTGGGTGGGGTGGAAACCAAGGCTTACCTCCACCGGTGT 240
Db      181 AAATTTGGGGGTATGCTTGTGGGTGGGGTGGAAACCAAGGCTTACCTCCACCGGTGT 240
Qy      241 GTTATTTGCTAACAGAGAGACATTTTCATGGGGCTAACAAAGAGACAAAGTTCAACAGCCTAT 300
Db      241 GTTATTTGCTAACAGAGAGAGACATTTTCATGGGGCTAACAAAGAGACAAAGTTCAACAGCCTAT 300
Qy      301 TACTTTGGCTCCCTCAGCCCAAGCTTCAGCTTATTCGATTTGATTCCTTCCAGGGAGAGAA 360
Db      301 TACTTTGGCTCCCTCAGCCCAAGCTTCAGCTTATTCGATTTGATTCCTTCCAGGGAGAGAA 360
Qy      361 ATCTATGCCCTTTCAGAGAGTCTGCTTCCATGCTTAAATCTCTGACGACATTTGTTGGG 420
Db      361 ATCTATGCCCTTTCAGAGAGTCTGCTTCCATGCTTAAATCTCTGACGACATTTGTTGGG 420
Qy      421 GGATGGGATATAGAGCAACATGAACCTGGCTGATGGCATGGCCAGAGGCAAAAGGTTTGAC 480
Db      421 GGATGGGATATAGAGCAACATGAACCTGGCTGATGGCATGGCCAGAGGCAAAAGGTTTGAC 480
Qy      481 ATCGATTTTGCAGAGAGAGTTGAGGCTTTACATGGAATCCATGGTTCCACTCCCGGAATC 540
Db      481 ATCGATTTTGCAGAGAGAGTTGAGGCTTTACATGGAATCCATGGTTCCACTCCCGGAATC 540
Qy      541 TACGACCCCGGATTTTCTCTGCTGCCAACCAAGAGAGCGTCCCAACAGCTGATTTAAGGCG 600
Db      541 TACGACCCCGGATTTTCTCTGCTGCCAACCAAGAGAGCGTCCCAACAGCTGATTTAAGGCG 600
Qy      601 ACAAAGCAAGAGCAAGTTTACGAAATCATCAAGACATCAAGACGCTTTAAGAGAGCACCC 660
Db      601 ACAAAGCAAGAGCAAGTTTACGAAATCATCAAGACATCAAGACGCTTTAAGAGAGCACCC 660
Qy      661 AAAGTGACAGAGGTGTTTCTCTGTGACTGCCCAACAGAGAGGTATAGCAATTTGGTT 720
Db      661 AAAGTGACAGAGGTGTTTCTCTGTGACTGCCCAACAGAGAGGTATAGCAATTTGGTT 720
Qy      721 GTAGGCTTTAATGACACCAATGAGAAATCTTGGCTGTGTGACAGAAATGAGGCTGAG 780
Db      721 GTAGGCTTTAATGACACCAATGAGAAATCTTGGCTGTGTGACAGAAATGAGGCTGAG 780
Qy      781 ATTTCTCTTCCACCTTGTATGCAATTTGCTGTGTGATGGAATAATGTTCTTTTCATTTAT 840
Db      781 ATTTCTCTTCCACCTTGTATGCAATTTGCTGTGTGATGGAATAATGTTCTTTTCATTTAT 840
Qy      841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGGCATGCGAGAAACACT 900
Db      841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGGCATGCGAGAAACACT 900
Qy      901 TTGATTTGTGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTTGTTGAT 960
Db      901 TTGATTTGTGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTTGTTGAT 960
Qy      961 TTTCTTGTGGGGCTGGTATCAAGCAACATCTATAGTTAGTTACAAACCATCTGGGAAC 1020
Db      961 TTTCTTGTGGGGCTGGTATCAAGCAACATCTATAGTTAGTTACAAACCATCTGGGAAC 1020
Qy      1021 AATGATGTATGATATCTCGGCTCCAAACCTTCGCTCCAAAGAAATCTTCCAGAGC 1080
Db      1021 AATGATGTATGATATCTCGGCTCCAAACCTTCGCTCCAAAGAAATCTTCCAGAGC 1080
Qy      1081 AAGCTTTGAGAGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTGAACATCCC 1140
Db      1081 AAGCTTTGAGAGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTGAACATCCC 1140

```

```

Qy      1141 GACCATGTTGTGTTAATTAAGTATGTGCTTACGTAGGGGATACAGAGAGCCATGAT 1200
Db      1141 GACCATGTTGTGTTAATTAAGTATGTGCTTACGTAGGGGATACAGAGAGCCATGAT 1200
Qy      1201 GAGTACATTTCAAGATATTTATGAGGTGGAAAGAACACATTTGTTTTCACAAACATGT 1260
Db      1201 GAGTACATTTCAAGATATTTATGAGGTGGAAAGAACACATTTGTTTTCACAAACATGT 1260
Qy      1261 GAGGATTCCTTTTAACTGCTCTATTTATCTTGAATGTTGTCCTTGTGCTGAGCTGAGC 1320
Db      1261 GAGGATTCCTTTTAACTGCTCTATTTATCTTGAATGTTGTCCTTGTGCTGAGCTGAGC 1320
Qy      1321 ACTAGAAATCAGTTTAAAGCTGAAATGAGGAAATTCATCTCATTCACCCAGTTGCT 1380
Db      1321 ACTAGAAATCAGTTTAAAGCTGAAATGAGGAAATTCATCTCATTCACCCAGTTGCT 1380
Qy      1381 ACCATTCTCAGCTATCTTACCAAGGCTCTCTGTTCCACGGGGTACACAGTGTGAT 1440
Db      1381 ACCATTCTCAGCTATCTTACCAAGGCTCTCTGTTCCACGGGGTACACAGTGTGAT 1440
Qy      1441 GCATTGTCAAGACGCTGCAATGCTGGAATAACATATGAGGGCTTGTGTGATTTGACC 1500
Db      1441 GCATTGTCAAGACGCTGCAATGCTGGAATAACATATGAGGGCTTGTGTGATTTGACC 1500
Qy      1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533
Db      1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533

```

```

RESULT 5
ADS82005
ID ADS82005 standard; cDNA; 1533 BP.
XX
AC ADS82005;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
  raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag=a
FT /product="myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
PI Hitz WD, Sebastian SA;
XX
DR WPI: 2004-63957/62.
DR P-PSDB; ADS82006.
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
  phosphate synthase, useful for altering raffinose saccharide, sucrose,
  phytic acid and inorganic phosphate content of soybean seeds.
XX
PS Claim 2; SEQ ID NO 15; 34pp; English.
XX

```

The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype with (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other:

Query Match 99.9%; Score 1531.4; DB 13; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGTCATGAGAAATTTAAAGTAGAGAGCTTAATGTAAGTACCCGAGCTGAGATT 60
DB 1 ATGTCATGAGAAATTTAAAGTAGAGAGCTTAATGTAAGTACCCGAGCTGAGATT 60
QY 61 CAGTCCGTGTACAACTACGAAACCAACCACTGTTCAGAGAACAGATGGACCTAT 120
DB 61 CAGTCCGTGTACAACTACGAAACCAACCACTGTTCAGAGAACAGATGGACCTAT 120
QY 121 CAGTGGATTGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTCCA 180
DB 121 CAGTGGATTGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTCCA 180
QY 181 AAATGGGGGTGATGCTTTGGGGTGGGGTGGAAACAAGGCTTACCCGTGAGT 240
DB 181 AAATGGGGGTGATGCTTTGGGGTGGGGTGGAAACAAGGCTTACCCGTGAGT 240
QY 241 GTTATGCTAACAGAGAGCAATTTGATGGGCTACAAAGAGCAAGATTCAACAGCCAT 300
DB 241 GTTATGCTAACAGAGAGCAATTTGATGGGCTACAAAGAGCAAGATTCAACAGCCAT 300
QY 301 TACTTTGCTCCCTCAACCAAGCTCAGCTATTGAGTTCGATCCTTCCAGGAGAGAA 360
DB 301 TACTTTGCTCCCTCAACCAAGCTCAGCTATTGAGTTCGATCCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCCATTCAGAGTCTGCTTCCATGCTTAATCTCGACCACTTGTGTTGGG 420
DB 361 ATCTATGCCCCATTCAGAGTCTGCTTCCATGCTTAATCTCGACCACTTGTGTTGGG 420
QY 421 GGATGGATATGCAAGCACTGAGTTCATGATGGCCAGGCAAGGAGGTTTGGAC 480
DB 421 GGATGGATATGCAAGCACTGAGTTCATGATGGCCAGGCAAGGAGGTTTGGAC 480
QY 481 ATCGATTTGCAAGAGCACTTGAAGCTTTACATGATTCATGGTTCCATCCCGGAAATC 540
DB 481 ATCGATTTGCAAGAGCACTTGAAGCTTTACATGATTCATGGTTCCATCCCGGAAATC 540

```

```

QY 541 TAGACCCGGATTTTCATGCTGCAACCAAGAGAGCGTCCAAACCAAGTAAAGGCG 600
DB 541 TAGACCCGGATTTTCATGCTGCAACCAAGAGAGCGTCCAAACCAAGTAAAGGCG 600
QY 601 ACAAAAGCAAGACCAAGTTCAGCAAAATCATCAAAAGACATCAAGGCTTTAAGAAAGCCACC 660
DB 601 ACAAAAGCAAGACCAAGTTCAGCAAAATCATCAAAAGACATCAAGGCTTTAAGAAAGCCACC 660
QY 661 AAAGTGACAAAGGTGTTCTCTGTGAGTCCCAACAGAGAGTATGCAATTTGGTT 720
DB 661 AAAGTGACAAAGGTGTTCTCTGTGAGTCCCAACAGAGAGTATGCAATTTGGTT 720
QY 721 GTAGGCTTAATGACCACTGAGATCTTGGGCTGTGAGCAAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTAATGACCACTGAGATCTTGGGCTGTGAGCAAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGATGATCCATGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTGATGATCCATGCTGTGATGAGAAATGTTCTTTCAATTAAT 840
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACCT 900
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACCT 900
QY 901 TTGATTGTGAGATGACCTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGAT 960
DB 901 TTGATTGTGAGATGACCTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGAT 960
QY 961 TTTCTTGTGGGGGCTGTATCAAGCAACATATATAGTTAGTTAAACCATCTGGGAAAC 1020
DB 961 TTTCTTGTGGGGGCTGTATCAAGCAACATATATAGTTAGTTAAACCATCTGGGAAAC 1020
QY 1021 AATGATGTATGATCTCTGCGCTCCAAACCTTCCGCTCCAGAGAAATCTCCAAAGC 1080
DB 1021 AATGATGTATGATCTCTGCGCTCCAAACCTTCCGCTCCAGAGAAATCTCCAAAGC 1080
QY 1081 AAGCTTTGTAAGATATGTTCAACAGCAATGCAATCTCTAAGACCTGTGTAACATCCC 1140
DB 1081 AAGCTTTGTAAGATATGTTCAACAGCAATGCAATCTCTAAGACCTGTGTAACATCCC 1140
QY 1141 GACCAATGTTGTTATTAATGATGCTTACGTAGGAGATAGCAAGAGCCATGAT 1200
DB 1141 GACCAATGTTGTTATTAATGATGCTTACGTAGGAGATAGCAAGAGCCATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATGTTTGCAACAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATGTTTGCAACAACATGT 1260
QY 1261 GAGGATTCCTTTAGCTGCTCTATTAATCTTGACCTTGCTCTTGCTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTAGCTGCTCTATTAATCTTGACCTTGCTCTTGCTGAGCTGAGC 1320
QY 1321 ACTAGAAATCAGTTTAAAGCTGAAATGAGGAAATTCATCTCAACCCAGTGTCT 1380
DB 1321 ACTAGAAATCAGTTTAAAGCTGAAATGAGGAAATTCATCTCAACCCAGTGTCT 1380
QY 1381 ACCATTCTCAGCTATCTGACCAAGGCTCCTGTTTCCACCGGATACCAAGTGTGAT 1440
DB 1381 ACCATTCTCAGCTATCTGACCAAGGCTCCTGTTTCCACCGGATACCAAGTGTGAT 1440
QY 1441 GCAATGTCAAGAGCACTGCTGCAATGAGGAGGCTTGTGTTGATTTGGCC 1500
DB 1441 GCAATGTCAAGAGCACTGCTGCAATGAGGAGGCTTGTGTTGATTTGGCC 1500
QY 1501 CCAGAGAAATACATGATCTCAGATCAAGTGA 1533
DB 1501 CCAGAGAAATACATGATCTCAGATCAAGTGA 1533

```

RESULT 6
ADS82003
ID ADS82003 standard; cdNA; 1533 BP.
XX

AC ADS82003;
XX
DT 18-NOV-2004 (first entry)
XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
XX
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; line 290180P03.
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
XX P-PSDB; ADS82004.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Example 8; SEQ ID NO 13; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.
XX
XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 99.9%; Score 1531.4; DB 13; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTCATCGAGAAATTTTAAAGTAGAGAGCTCTATGTAAGTACACGAGACTGAGATT 60
DB 1 ATGTCATCGAGAAATTTTAAAGTAGAGAGCTCTATGTAAGTACACGAGACTGAGATT 60
QY 61 CAGTCGCGTGTACAACTACGAAACCCAGCACTTGTTCAGAAACGGAATGCGACTAT 120
DB 61 CAGTCGCGTGTACAACTACGAAACCCAGCACTTGTTCAGAAACGGAATGCGACTAT 120
QY 121 CAGTGGATTGTCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCACTGTTCA 180
DB 121 CAGTGGATTGTCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCACTGTTCA 180
QY 181 AAATGGGGGGTGTATGCTGTGGGTTGGGGTGAACCAACGGCTTACCTCAACGGTGT 240
DB 181 AAATGGGGGGTGTATGCTGTGGGTTGGGGTGAACCAACGGCTTACCTCAACGGTGT 240
QY 241 GTTATTGCTTACAGAGAGGACATTTTCATGGGCTACAAAGAACATTTCAACAGCCAA 300
DB 241 GTTATTGCTTACAGAGAGGACATTTTCATGGGCTACAAAGAACATTTCAACAGCCAA 300
QY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCGAGTTGGATCCTTCCAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCGAGTTGGATCCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCAATTCAGAGTCTGCTTCAATGTTAAATCTGACAGCATTTGTGTGG 420
DB 361 ATCTATGCCCAATTCAGAGTCTGCTTCAATGTTAAATCTGACAGCATTTGTGTGG 420
QY 421 GGATGGATATCAGCAAGATGAACTTGGCTGATGTCATGCGGCAAGAGTGTGTTGAC 480
DB 421 GGATGGATATCAGCAAGATGAACTTGGCTGATGTCATGCGGCAAGAGTGTGTTGAC 480
QY 481 ATCGATTGTCAGAGCAGTTGAGGCTTACATGAAATCATGTTCCACTCCCGGAATC 540
DB 481 ATCGATTGTCAGAGCAGTTGAGGCTTACATGAAATCATGTTCCACTCCCGGAATC 540
QY 541 TAGGACCCGGAATTTTCATGCTGCTCCAAACCAAGAGAGCGTCCAAACAGTATTAAGGC 600
DB 541 TAGGACCCGGAATTTTCATGCTGCTCCAAACCAAGAGAGCGTCCAAACAGTATTAAGGC 600
QY 601 ACAAAGCAAGACCAAGTTCAGCAAAATCAACAAAGCATCAAGCGTTTAAAGAACCC 660
DB 601 ACAAAGCAAGACCAAGTTCAGCAAAATCAACAAAGCATCAAGCGTTTAAAGAACCC 660
QY 661 AAAGTGACAGAGTGTGTTCTCTGTGACTGCCAACAGAGAGTATAGCAATTTGGTT 720
DB 661 AAAGTGACAGAGTGTGTTCTCTGTGACTGCCAACAGAGAGTATAGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTGGGCTGTGAGAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTTGGGCTGTGAGAGAAATGAGGCTGAG 780
QY 781 ATTTCCTCTTCCACCTTGTATGCCATTGCTGTGATGGAATAATGTTCTTCAATTA 840
DB 781 ATTTCCTCTTCCACCTTGTATGCCATTGCTGTGATGGAATAATGTTCTTCAATTA 840
QY 841 GGAAGCCCTCAAGAACTTTTGTACAGGGCTGATTTGATCTTGGCAATGCGAGAACT 900
DB 841 GGAAGCCCTCAAGAACTTTTGTACAGGGCTGATTTGATCTTGGCAATGCGAGAACT 900
QY 901 TTGATTGGTGAAGTACTTCAAGAGTGTGAGCAACAAATGAATCTGTGTGTTGAT 960
DB 901 TTGATTGGTGAAGTACTTCAAGAGTGTGAGCAACAAATGAATCTGTGTGTTGAT 960
QY 961 TTTCTTGTGGGGCTGTGTATCAAGCCACATCATATAGTTATCAACCATCTGGGAAAC 1020
DB 961 TTTCTTGTGGGGCTGTGTATCAAGCCACATCATATAGTTATCAACCATCTGGGAAAC 1020
QY 1021 AATGATGTATGAAATCTCGGCTCCAAACCTTCGCTCCAGGAATCTTCCAGAGC 1080
DB 1021 AATGATGTATGAAATCTCGGCTCCAAACCTTCGCTCCAGGAATCTTCCAGAGC 1080

```

Db 1021 AATGATGTATGATCTCTGGGCTCCAAACCTTCGGTCCAAAGAAATCTCCAGAGC 1080
QY 1081 AACGTTGTGACATATGTCACACCAATGCCATCTCTATATAGCTGTGAACTCC 1140
Db 1081 AACGTTGTGACATATGTCACACCAATGCCATCTCTATATAGCTGTGAACTCC 1140
QY 1141 GACCATGTGTTGTATTAAGTATGTCCTTACGTAAGGGGATAGCAAGAGCCATGGAT 1200
Db 1141 GACCATGTGTTGTATTAAGTATGTCCTTACGTAAGGGGATAGCAAGAGCCATGGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATTTGTTGCAACAACATGT 1260
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATTTGTTGCAACAACATGT 1260
QY 1261 GAGGATTCCTTTTACCTGCTCTATATATCTTATGACTTGGTCTTTTGTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTTACCTGCTCTATATATCTTATGACTTGGTCTTTTGTGAGCTGAGC 1320
QY 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGGGAAATTCGACTCATTTCCACCGAGTTGCT 1380
Db 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGGGAAATTCGACTCATTTCCACCGAGTTGCT 1380
QY 1381 ACCATTCCTCAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGTGAAT 1440
Db 1381 ACCATTCCTCAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGTGAAT 1440
QY 1441 GCAATGTCAAACGAGCTGCAATGCTGGAAGAAATATGAGGGCTTGTGTGATTTGGCC 1500
Db 1441 GCAATGTCAAACGAGCTGCAATGCTGGAAGAAATATGAGGGCTTGTGTGATTTGGCC 1500
QY 1501 CCAGAGAAATACATGATCTCTCGAGTACAGTGA 1533
Db 1501 CCAGAGAAATACATGATCTCTCGAGTACAGTGA 1533

```

RESULT 7
AD014498
ID AD014498 standard; cDNA; 1533 BP.

AC AD014498;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.

KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.

OS Glycine max.

XX Synthetic.

Key Location/Qualifiers
1. 1533

FT CDS /*tag= a
FT /product= "mutant soybean myo-inositol 1-phosphate
FT synthase #2"

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718955.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 26-APR-1999; 99US-00299315.

XX 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.

```

XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
PI MPI; 2004-533135/51.
XX P-PSDB; AD014499.
DR
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Example 8; SEQ ID NO 9; 48bp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

```

Query Match 95.5%; Score 1464.2; DB 12; Length 1533;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

QY 1 ATGTTCAATCAGAAATTTTAAGGTAGAGAGTCTTAATGTGAATGACACCGAGACTGAGATT 60
Db 1 ATGTTCAATCAGAAATTTTAAGGTAGAGAGTCTTAATGTGAATGACACCGAGACTGAGATT 60
QY 61 CAGTCGGGTGTAACAACTAGCAACCAACCACTTTTCCAGAAACAGAAATGGCACTTAT 120
Db 61 CAGTCGGGTGTAACAACTAGCAACCAACCACTTTTCCAGAAACAGAAATGGCACTTAT 120
QY 121 CAGTGAATTTGCAACCCAAATCGTCAACTGCAATTTAAACCAACCAACCATGTTCCA 180
Db 121 CAGTGAATTTGCAACCCAAATCGTCAACTGCAATTTAAACCAACCAACCATGTTCTT 180
QY 181 AAATTTGGGGGTGATGCTTTGGGTTGGGGTGGAAACAACGCTCTTACCTCAACCGGTGT 240
Db 181 AAATTTAGGGGTATATGCTTTGGGTTGGGGTGGAAACAACGCTCTTACCTCAACCGGTGT 240
QY 241 GTTATTTGCTAACAGAGAGACATTTTCAATGAGGCTTACAAAGACACAAATTCACACCAAT 300
Db 241 GTTATTTGCTAACCGAGAGGCAATTTTCAATGAGGCTTACAAAGACACAAATTCACACCAAT 300
QY 301 TACTTTGGCTCCCTACCAAGCTCAGCTATTGAGTTGATTCCTTCCAGGGAGAGGAA 360
Db 301 TACTTTGGCTCCCTACCAAGCTCAGCTATTGAGTTGATTCCTTCCAGGGAGAGGAA 360
QY 361 ATCTATGCCCCATTCAGAGCTGCTTCAATGTTAACTCCAGATTTGTTGGG 420
Db 361 ATCTATGCCCCATTCAGAGCTGCTTCAATGTTAACTCCAGATTTGTTGGG 420
QY 421 GGAATGGGATATCAGAAATATGAACTGCTGATGCAATGAGGCAAGGATTTTAC 480
Db 421 GGAATGGGATATCAGAAATATGAACTGCTGATGCAATGAGGCAAGGATTTTAC 480
QY 481 ATGCAATTTGCAAGAGCAATGAGGCTTATCAGTGAATCATGATGTTCCACTCCCGGAATC 540
Db 481 ATGCAATTTGCAAGAGCAATGAGGCTTATCAGTGAATCATGATGTTCCACTCCCGGAATC 540

```

```
QY 541 TACGACCCCGATTTCATTGCTGCTCCAA CCAAGAGACGTCACCAACACCTGATTAAAGGC 600
DB 541 TATGACCCCGATTTCATTGCTGCTCCAA CCAAGAGACGTCACCAACACCTGATTAAAGGC 600
QY 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAAAGACATCAAGCCGTTTAAAGAACCCACC 660
DB 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAAAGACATCAAGCCGTTTAAAGAACCCACC 660
QY 661 AAAGTGACAAAGTGGTGTCTCGTGAGACTGCGCAACAGAGAGGATATGCAATTTGGTT 720
DB 661 AAAGTGACAAAGTGGTGTCTCGTGAGACTGCGCAACAGAGAGGATATGCAATTTGGTT 720
QY 721 GTAGGCTTAAATGACACCAATGAGATCTCTTGCTGCTGTGAGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTAAATGACACCAATGAGATCTCTTGCTGCTGTGAGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGTATGCGCATGCGCTGTGATGAGAAATGTTCTTCTTCAATAT 840
DB 781 ATTTCTCTTCCACCTTGTATGCGCATGCGCTGTGATGAGAAATGTTCTTCTTCAATAT 840
QY 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTGATCTTGCCATCGCGAGAACACT 900
DB 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTGATCTTGCCATCGCGAGAACACT 900
QY 901 TTGATTTGGTGAAGATGACTTCAAGAGTGGTCAAGCCAAATGAAATCTGTGTGTTGAT 960
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGGTCAAGCCAAATGAAATCTGTGTGTTGAT 960
QY 961 TTTCTTTGGGGGGCTGTGATCAAGCCAAATCATATAGTTAAGTCAACCATCTGGGAAC 1020
DB 961 TTTCTTTGGGGGGCTGTGATCAAGCCAAATCATATAGTTAAGTCAACCATCTGGGAAC 1020
QY 1021 AATGATGATGATGATCTCTCGGCTCCCAACACCTTCCGCTCCAGAGAAATCTCCAAGAC 1080
DB 1021 AATGATGATGATGATCTCTCGGCTCCCAACACCTTCCGCTCCAGAGAAATCTCCAAGAC 1080
QY 1081 AACGTTGTTGACGATATGCTCAACAGCAATGCGCATCTCTATGAGCCTTGATCAATCCC 1140
DB 1081 AACGTTGTTGACGATATGCTCAACAGCAATGCGCATCTCTATGAGCCTTGATCAATCCC 1140
QY 1141 GACCATGTTGTTGTTATTAAGTATGCTCTTACGTAAGGAGATAGCAAGAGACCATGAT 1200
DB 1141 GACCATGTTGTTGTTATTAAGTATGCTCTTACGTAAGGAGATAGCAAGAGACCATGAT 1200
QY 1201 GAGTACCTTCAAGATATTCATGAGTGGTGAAGAACACCATGTTTGGACAACACATGT 1260
DB 1201 GAGTACCTTCAAGATATTCATGAGTGGTGAAGAACACCATGTTTGGACAACACATGT 1260
QY 1261 GAGGATTTCCCTTTTAACTGCTCTATATCTTGAATGCTGCTTCTTGGCTGAGCTGAGC 1320
DB 1261 GAGGATTTCCCTTTTAACTGCTCTATATCTTGAATGCTGCTTCTTGGCTGAGCTGAGC 1320
QY 1321 ACTAGATTCAGTTTAAAGTGAAGTGAAGGAAATTCACACTCATTCCACCCAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGTGAAGTGAAGGAAATTCACACTCATTCCACCCAGTTGCT 1380
QY 1381 ACCATCTCAGCATGACCAAGAGTCCCTGCTGTTCCACCGGATCACCAAGTGGTGAAT 1440
DB 1381 ACCATCTCAGCATGACCAAGAGTCCCTGCTGTTCCACCGGATCACCAAGTGGTGAAT 1440
QY 1441 GGAATGTCAAAGACGCTGCAATGCTGAAACATTAATGAGGCTTGTGTTGATTGGCC 1500
DB 1441 GGAATGTCAAAGACGCTGCAATGCTGAAACATTAATGAGGCTTGTGTTGATTGGCC 1500
QY 1501 CCAGAGAAATTAACATGATTTCTCGAGTCAAGTGA 1533
DB 1501 CCAGAGAAATTAACATGATTTCTCGAGTCAAGTGA 1533
```

RESULT 8
ADS81999
ID ADS81999 standard; cDNA, 1533 BP.

```
XX AC ADS81999;
XX XX 18-NOV-2004 (first entry)
DT DT
XX XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX XX
OS Glycine max; line 29004JP01.
XX XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag=a
FT /product="myo-inositol 1-phosphate synthase"
XX XX
PN US2003074685-A1.
XX XX
XX 17-APR-2003.
XX XX
PF 11-MAR-2002; 2002US-00025003.
XX XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US0006822.
XX XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX XX
PI Hitz WD, Sebastian SA;
XX XX
DR WPI; 2004-639957/62.
DR P-PsDB; ADS82000.
XX XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX XX
PS Example 8; SEQ ID NO 9; 34p; English.
XX XX
CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphatase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.
XX XX
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
```

Query Match 95.5%; Score 1464.2; DB 13; Length 1533;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1450; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

QY 1 ATGTCATCGAGAAATTTTAAGTAGAGAGCTCTAATGTGAAATGACCCGAGACTGAGATT 60
DB 1 ATGTCATCGAGAAATTTTAAGTAGAGAGCTCTAATGTGAAATGACCCGAGACTGAGATT 60
QY 61 CAGTCCGTGTACACATACGAAACACCGAACTTGTTCAGAGAACGAAATGCGACTTAT 120
DB 61 CAGTCCGTGTACACATACGAAACACCGAACTTGTTCAGAGAACGAAATGCGACTTAT 120
QY 121 CAGTGAATGTCAAAACCAATCGTCACTACCAATTTAAACCAACCCATGTTTCA 180
DB 121 CAGTGAATGTCAAAACCAATCGTCACTACCAATTTAAACCAACCCATGTTTCT 180
QY 181 AAATTGGGGGTGATGCTTGTGGGTGGGTGAAACACGCTCTACCTCAGCGTGGT 240
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGTGAAACACGCTCTACCTCAGCGTGGT 240
QY 241 GTTATTGCTAACAGAGACATTTCAATGGGCTAACAGAACAGATTCAACAGCCAAAT 300
DB 241 GTTATTGCTAACAGAGACATTTCAATGGGCTAACAGAACAGATTCAACAGCCAAAT 300
QY 301 TACTTTGGCTCCCTCAACCCAGCTGAGCTATCGAGTTGGGCTCTTCCAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCCAGCTGAGCTATCGAGTTGGGCTCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCATTTCAAGAGTCTGCTTCAATGTTTATCTGACGACATTTGTTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCAATGTTTATCTGACGACATTTGTTTGGG 420
QY 421 GGAATGGATATACGACAACTGAACCTTGCTGATGCAATGGCCAGGCAAAAGTGTGAC 480
DB 421 GGAATGGATATACGACAACTGAACCTTGCTGATGCAATGGCCAGGCAAAAGTGTGAC 480
QY 481 ATCGAATTTGACAGAGAGTTGAGGCTTTCATGGAATTCATGCTTCCACCTCCCGGAATC 540
DB 481 ATCGAATTTGACAGAGAGTTGAGGCTTTCATGGAATTCATGCTTCCACCTCCCGGAATC 540
QY 541 TACGACCCCGAATTTCAATGCTGCAACCAAGAGAGCTGCCAACACGTTAATAGGCG 600
DB 541 TACGACCCCGAATTTCAATGCTGCTGCAACCAAGAGAGCTGCCAACACGTTAATAGGCG 600
QY 601 ACAAGACAGAGCAAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAACCCACC 660
DB 601 ACAAGACAGAGCAAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAACCCACC 660
QY 661 AAAGTGGACAGAGTGTGCTGCTGAGACCTGCCAACACAGAGGTAATGCAATTTGGTT 720
DB 661 AAAGTGGACAGAGTGTGCTGCTGAGACCTGCCAACACAGAGGTAATGCAATTTGGTT 720
QY 721 GTAGGCTTAAATGACACATGAGAACTCTTGGCTGCTGTGACAGAAATGAGCTGAG 780
DB 721 GTAGGCTTAAATGACACATGAGAACTCTTGGCTGCTGTGACAGAAATGAGCTGAG 780
QY 781 AATTTCTCTTCACTTGTATGCAATGCTGTGTGATGAAATGTTCTTTCAATTAAT 840
DB 781 AATTTCTCTTCACTTGTATGCAATGCTGTGTGATGAAATGTTCTTTCAATTAAT 840
QY 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTGGCATCGCGAGAACACT 900
DB 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTGGCATCGCGAGAACACT 900
QY 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGACCAAAATGAAATCTGTGTTGGTAT 960
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGACCAAAATGAAATCTGTGTTGGTAT 960
QY 961 TTTCTTTGGGGGCTGTGATCAAGCCACATCTAATAGTTTCAACCATCTGGGAAC 1020
DB 961 TTTCTTTGGGGGCTGTGATCAAGCCACATCTAATAGTTTCAACCATCTGGGAAC 1020
QY 1021 AATGATGATATGATCTCTCGGCTCCACAAACCTTCGCTCCAGGAATCTTCCAGAGC 1080
  
```

```

DB 1021 AATGATGATATGATCTTTCGGCTCCACAAACCTTTCGTTCCAGGAATCTTCCAGAGC 1080
QY 1081 AACGTTGTGACGATATGCTCAACAGCAATGCAATCTCTATGAGCTGTGGAAATCCC 1140
DB 1081 AACGTTGTGATGATATGCTCAACAGCAATGCAATCTCTATGAGCTGTGGAAATCCA 1140
QY 1141 GACCATGTTGTTTATTAAGTATGCTTACGTAAGGGAATGACAAAGAGCCATGAT 1200
DB 1141 GACCATGTTGTTTATTAAGTATGCTTACGTAAGGGAACAGCAAGAGCCATGAT 1200
QY 1201 GAGTACCTTCAGAGATATTCATGCGGTGAAAGAACCAATGTTTGTGACACACATGT 1260
DB 1201 GAGTACCTTCAGAGATATTCATGCGGTGAAAGAGCAATGTTTGTGACACACATGC 1260
QY 1261 GAGGATTCCTTTTACGCTGCTCTATTAATCTTGAATGATCTTCTGAGCTGAGC 1320
DB 1261 GAGGATTCCTCTCTTACGCTGCTCTATTAATCTTGAATGATGATCTTCTGAGCTCAGC 1320
QY 1321 ACTAGATTCAGTTTAAAGTGAATAAGGGAATTCACATCTATTCCACCGATTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGTGAATAAGGGAATTCACATCTATTCCACCGATTGCT 1380
QY 1381 ACCATTCAGCTATCTGACCAAGGCTCCTGTTTCCACCGGATACACAGTGTGAAT 1440
DB 1381 ACCATTCAGCTATCTGACCAAGGCTCCTGTTTCCACCGGATACACAGTGTGAAT 1440
QY 1441 GCATTTGCAAGCAGCTGCAATGCTGAAACCATTAATGAGGCTTGTGATGGCC 1500
DB 1441 GCATTTGCAAGCAGCTGCAATGCTGAAACCATTAATGAGGCTTGTGATGGCC 1500
QY 1501 CCAGAGATTAACATGATTTCTGAGTACAGTGA 1533
DB 1501 CCAGAGATTAACATGATTTCTGAGTACAGTGA 1533

```

RESULT 9
 ADS81993
 ID ADS81993 standard; cDNA; 1533 BP.

AC ADS81993;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
 XX
 KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 OS Glycine max; cultivar Wye.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"

PN US2003074685-A1.
 PD 17-APR-2003.
 XX
 PF 11-MAR-2002; 2002US-00025003.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 XX
 PI Hitz WD, Sebastian SA;
 XX
 DR MPI; 2004-639957/62.
 DR P-PsDB; ADS81994.

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX Claim 2; SEQ ID NO 1; 34bp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

XX Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Query Match 95.5%; Score 1464.2; DB 13; Length 1533;

Beet Local Similarity 97.2%; Pred. No. 0; Mismatches 43; Indels 0; Gaps 0;

Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 ATGTCATCGAGAAATTTAAGGTAGAGAGTCTTAATGTGAATGACACGAGACTGAGATT 60
DB 1 ATGTCATCGAGAAATTTAAGGTAGAGAGTCTTAATGTGAATGACACGAGACTGAGATT 60
QY 61 CAGTCCGTGTACATGACGAAACACCGAATTGTTTCAAGAGACGGAATGCGACTAT 120
DB 61 CAGTCCGTGTACATGACGAAACACCGAATTGTTTCAAGAGACGGAATGCGACTAT 120
QY 121 CAGTGGATGTGAACCCAAATCCGTCATCCATTTAAACCAACGCCATGTTTCA 180
DB 121 CAGTGGATGTGAACCCAAATCCGTCATCCATTTAAACCAACGCCATGTTTCT 180
QY 181 AATTTGGGGGTGATGCTTGTGGGTGGGAAACACGCGCTTCAACCGGTGAT 240
DB 181 AATTTGGGGGTGATGCTTGTGGGTGGGAAACACGCGCTTCAACCGGTGAT 240
QY 241 AATTTGGGGGTGATGCTTGTGGGTGGGAAACACGCGCTTCAACCGGTGAT 240
DB 241 AATTTGGGGGTGATGCTTGTGGGTGGGAAACACGCGCTTCAACCGGTGAT 240
QY 241 GTTATTTGCTAACGAGAGACATTTTATGCGGCTTCAACGAGACATGTTTCAACGAGCAAT 300
DB 241 GTTATTTGCTAACGAGAGACATTTTATGCGGCTTCAACGAGACATGTTTCAACGAGCAAT 300
QY 301 TACTTTGGGCTCCCTCCACCAAGCTCAGCTATTCGAGTTGGGCTTCCAGGAGAGGAA 360
DB 301 TACTTTGGGCTCCCTCCACCAAGCTCAGCTATTCGAGTTGGGCTTCCAGGAGAGGAA 360
QY 361 ATCTATGCCCATTTCAAGAGCTGCTTCAATGTTAATCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGCTGCTTCAATGTTAATCTGACGACATTTGTTGGG 420
QY 421 GGATGGATATGACGACATGAACTGTGCTGATGCTATGCGCAAGGCAAGGTGTTGAC 480
DB 421 GGATGGATATGACGACATGAACTGTGCTGATGCTATGCGCAAGGCAAGGTGTTGAC 480

DB 421 GGATGGATATGACGACATGAACTGTGCTGATGCTATGCGCAAGGCAAGGTGTTGAC 480
QY 481 ATCTATGCCCATTTCAAGAGCTGCTTCAATGTTAATCTGACGACATTTGTTGGG 540
DB 481 ATCTATGCCCATTTCAAGAGCTGCTTCAATGTTAATCTGACGACATTTGTTGGG 540
QY 541 TACGACCCGGATTTTATGCTGCGCAACCAAGAGAGCTGCCAACACAGTATTAAAGGC 600
DB 541 TATGACCCGGATTTTATGCTGCGCAACCAAGAGAGCTGCCAACACAGTATTAAAGGC 600
QY 601 ACAAAGCAAGGCAAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAGCCACC 660
DB 601 ACAAAGCAAGGCAAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAGCCACC 660
QY 661 AAGGTGCAAGAGTGTGCTGCTGAGCTGCAACACAGAGAGTATGCAATTTGGTT 720
DB 661 AAGGTGCAAGAGTGTGCTGCTGAGCTGCAACACAGAGAGTATGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACTTGTATGCAATGCGCTGCTGATGAGAAATGTTCTTTCTTATAT 840
DB 781 ATTTCTCTTCCACTTGTATGCAATGCGCTGCTGATGAGAAATGTTCTTTCTTATAT 840
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATCGCAGGAACCT 900
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATCGCAGGAACCT 900
QY 901 TTGATTTGGTGAATGATCTTCAAGAGTGTGCAAGCAAAATGAAATCTGTGTGTTGAT 960
DB 901 TTGATTTGGTGAATGATCTTCAAGAGTGTGCAAGCAAAATGAAATCTGTGTGTTGAT 960
QY 961 TTTCTTGGGGGGCGTGTATGACGCAACATCTATAGTTAGTTCAACATCTGGGAAC 1020
DB 961 TTTCTTGGGGGGCGTGTATGACGCAACATCTATAGTTAGTTCAACATCTGGGAAC 1020
QY 1021 AATGATGATGAATCTCTCGGCTTCACAAACCTTCCGCTCAGGAAGATCTCAAGAGC 1080
DB 1021 AATGATGATGAATCTCTCGGCTTCACAAACCTTCCGCTCAGGAAGATCTCAAGAGC 1080
QY 1081 AAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 AAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GACCAATGTTGTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GACCAATGTTGTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTTCAAGGCTGGAAGAAACACATTTGTTTTCACAAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTTCAAGGCTGGAAGAAACACATTTGTTTTCACAAACATGT 1260
QY 1261 GAGATTCCTTTTATGCTTCTTATTTATTTTGAATTTTGTGCTTCTTGTGAGCTGAGC 1320
DB 1261 GAGATTCCTTTTATGCTTCTTATTTTATTTTGAATTTTGTGCTTCTTGTGAGCTGAGC 1320
QY 1321 ACTAAGATTCAGTTTAAAGCTGGAATGAGGAAATTCCTCACTCATCCCACTTGTCT 1380
DB 1321 ACTAAGATTCAGTTTAAAGCTGGAATGAGGAAATTCCTCACTCATCCCACTTGTCT 1380
QY 1381 ACCATTTCTCAGCTATCTGACCAAGGCTCTCTGTTTTCACCGGCTTACCAAGTGTGAT 1440
DB 1381 ACCATTTCTCAGCTATCTGACCAAGGCTCTCTGTTTTCACCGGCTTACCAAGTGTGAT 1440
QY 1441 GCATTTGCAAGCAGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 1500
DB 1441 GCATTTGCAAGCAGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 1500
QY 1501 CCAAGGAATTAACATGATTTCTGAGTAAAGTGA 1533
DB 1501 CCAAGGAATTAACATGATTTCTGAGTAAAGTGA 1533

RESULT 10
ID ADQ14490 standard; cDNA; 1760 BP.
AC ADQ14490;
XX 23-SEP-2004 (first entry)
DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #1.
XX
KM Soybean; myo-inositol 1-phosphate synthase; gene; ss:
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KM inorganic phosphate.
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 54..1586
FT /tag= a
FT /product= "Wild type soybean myo-inositol 1-phosphate
FT synthase #1"
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
DR WPI: 2004-533135/51.
XX
XX P-PSDB; ADQ14491.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Claim 4; SEQ ID NO 1; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Query Match 95.5%; Score 1464.2; DB 12; Length 1760;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 1 ATGTCATGAGAAATTTTAAAGTAGAGAGCTTATGTGAATGACCCGAGCTGAGATT 60
DB 54 ATGTCATGAGAAATTTTAAAGTAGAGAGCTTATGTGAATGACCCGAGCTGAGATT 113
QY 61 CAGTCCGTGTACAACTACAGAAACCAACGAACTTGTTCAGAGAAACAGAAATGACCTAT 120
DB 114 CAGTCCGTGTACAACTACAGAAACCAACGAACTTGTTCAGAGAAACAGAAATGACCTAT 173
QY 121 CAGTGAATGTCAAAATCCCAATTCCTCAACTACAAATTTTAAACCAACCACTGTTCCA 180
DB 174 CAGTGAATGTCAAAATCCCAATTCCTCAACTACAAATTTTAAACCAACCACTGTTCC 233
QY 181 AAATGGGGGTGATGCTGTGGGGGTGGGAAACAGGCTTCCTCCACCGGTGT 240
DB 234 AAATGGGGGTGATGCTGTGGGGGTGGGAAACAGGCTTCCTCCACCGGTGT 293
QY 241 GTTATTTCTAACAAGAGAGCAATTTTCATGAGGCTTACAAAGACAAAGATTAACAGCAAT 300
DB 294 GTTATTTCTAACAAGAGAGCAATTTTCATGAGGCTTACAAAGACAAAGATTAACAGCAAT 353
QY 301 TACTTTGGCTCCTTCAACCAAGCTTACGTAATTCGAGTTGATCCTTCCAGGAGAGAA 360
DB 354 TACTTTGGCTCCTTCAACCAAGCTTACGTAATTCGAGTTGATCCTTCCAGGAGAGAA 413
QY 361 ATCTATGCCCCATTCAAGAGTCTGCTTCAATGTTTAACTCGAAGCAATTTGTTGG 420
DB 414 ATCTATGCCCCATTCAAGAGTCTGCTTCAATGTTTAACTCGAAGCAATTTGTTGG 473
QY 421 GGATGGGATATGACCAATGAACTGCTGATGTCATGAGGCGCAAGGAGGATTTGAC 480
DB 474 GGATGGGATATGACCAATGAACTGCTGATGTCATGAGGCGCAAGGAGGATTTGAC 533
QY 481 ATGATTTGCAAGAGAGTTGAGGCTTATCATGAAATTCATGTTCCCTCCCGAATC 540
DB 534 ATGATTTGCAAGAGAGTTGAGGCTTATCATGAAATTCATGTTCCCTCCCGAATC 593
QY 541 TACGACCCGGATTTTCAATGCTGCTCCCAACCAAGAGGCGGCAACCAACCTGTTAAGGGC 600
DB 594 TACGACCCGGATTTTCAATGCTGCTCCCAACCAAGAGGCGGCAACCAACCTGTTAAGGGC 653
QY 601 ACAAAACAAGACAAGTTGACGAAATCATCAAGAATCAAGAGGCTTTAAGAGGACACC 660
DB 654 ACAAAACAAGACAAGTTGACGAAATCATCAAGAATCAAGAGGCTTTAAGAGGACACC 713
QY 661 AAAGTGACAGAGTGTGCTCTGTGAGCTGCCAACACAGAGGTTAAGCAATTTGGTT 720
DB 714 AAAGTGACAGAGTGTGCTCTGTGAGCTGCCAACACAGAGGTTAAGCAATTTGGTT 773
QY 721 GTAGGCTTAAAGACACCATGAGAAATCTTGGCTGCTGCTGACAGAAATGAGCTGAG 780
DB 774 GTAGGCTTAAAGACACCATGAGAAATCTTGGCTGCTGCTGACAGAAATGAGCTGAG 833
QY 781 ATTTCCTCTTCAACCTTGTATGCTATGCTGTGTGATGAGAAATGTTCTTCAATTAAT 840
DB 834 ATTTCCTCTTCAACCTTGTATGCTATGCTGTGTGATGAGAAATGTTCTTCAATTAAT 893
QY 841 GGAAGCCCTCAGACACCTTTGTACCAAGGCTGATTTGATCTTGGCAATGCGAGAACT 900
DB 894 GGAAGCCCTCAGACACCTTTGTACCAAGGCTGATTTGATCTTGGCAATGCGAGAACT 953
QY 901 TTGATTTGAGAGAGACTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGTAT 960
DB 954 TTGATTTGAGAGAGACTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGTAT 1013
QY 961 TTTCTTGTGGGGCTGTGTATCAAGCAACATCTATAGTTAGTTACCAACATCTGGGAAAC 1020
DB 1014 TTTCTTGTGGGGCTGTGTATCAAGCAACATCTATAGTTAGTTACCAACATCTGGGAAAC 1073
QY 1021 AATGATGATGATCTCTGGGCTCCAAACCTTCCGCTCCAAAGAAATCTCCAAAGAC 1080

```
Db 1074 AATGATGATGATGATCTTCGGCTCCAACTTCCGTTCCAGAAATCTCCAAGAGC 1133
Oy 1081 AACGTTGTTGAGATATGTTCAACAGCAATGCCATCCTTATAGCCTGTTAAATGCC 1140
Db 1134 AACGTTGTTGAGATATGTTCAACAGCAATGCCATCCTTATAGCCTGTTAAATGCC 1193
Oy 1141 GACCATGTTGTTATTAATGATGCTTACGTAAGGGATAGCAAGAGCCATGAT 1200
Db 1194 GACCATGTTGTTATTAATGATGCTTACGTAAGGGATAGCAAGAGCCATGAT 1253
Oy 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACCATTTGTTTGCAACATGCT 1260
Db 1254 GAGTACACTTCAAGATATTCATGGGTGAAAGAACCATTTGTTTGCAACATGCT 1313
Oy 1261 GAGGATTCCTTTTGGCTGCTCTTATATCTGAGCTGGTCTCTTGGCTGAGTGAGC 1320
Db 1314 GAGGATTCCTTTTGGCTGCTCTTATATCTGAGCTGGTCTCTTGGCTGAGTGAGC 1373
Oy 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCATCTCATTCACCCAGTTGCT 1380
Db 1374 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCATCTCATTCACCCAGTTGCT 1433
Oy 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTGTTTCCACCGGTAACACCAAGTGTGAT 1440
Db 1434 ACCATTTCTAGCTATCTGACCAAGGCTCTGTTTCCACCGGTAACACCAAGTGTGAT 1493
Oy 1441 GATTTGTCAAACAGAGCTGCAATGCTGAAACATTAATAGGGCTGTTGTTGATTTGGCC 1500
Db 1494 GATTTGTCAAACAGAGCTGCAATGCTGAAACATTAATAGGGCTGTTGTTGATTTGGCC 1553
Oy 1501 CCAGAGATTAACATGATTTCTCGAGTACAGTGA 1533
Db 1554 CCAGAGATTAACATGATTTCTCGAGTACAGTGA 1586
```

RESULT 11

AAV62440
ID AAV62440 standard; cDNA; 1782 BP.

AAV62440;

17-OCT-2003 (revised)
02-FEB-1999 (first entry)

Soybean wild-type myo-inositol 1-phosphate synthase cDNA.

KW Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW phytic acid; ds.

OS Glycine max; line LR13.

PH Key Location/Qualifiers
FT CDS 54..1586
/*tag= a

W09845448-A1.

15-OCT-1998.

07-APR-1998; 98WO-US006822.

08-APR-1997; 97US-00835751.

(DUPO) DU PONT DE NEMOURS & CO E I.

Hitz WD, Sebastian SA;

WPI; 1998-568353/48.

P-PSDB; AAW79740.

Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,

PT phytic acid, etc.

PS Example 5; Page 44-45; 63pp; English.

XX This is the nucleotide sequence of cDNA encoding the wild-type soybean
XX myo-inositol 1-phosphate synthase (MI 1-PS) present in clone p5mi-1ps
CC (ATCC 97970). The clone was isolated from a cDNA library of soybean line
CC LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
CC thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,
CC raffinose and stachyose. A mutant MI 1-PS nucleic acid (see AAV62443) has
CC been identified in soybean line LR13, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC (G to T at base 1241) in the LR13 sequence. The mutation results in a
CC seed phenotype of very low raffinose saccharide sugars, very high sucrose
CC and low phytic acid. The nucleic acid is used to alter the raffinose
CC saccharide, sucrose, phytic acid and inorganic phosphate content of
CC soybean seeds, leading to useful soybean products, e.g. a seed phytic
CC acid content of less than 17 ug/g, a seed content of raffinose and
CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC greater than 200 ug/g. (updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1782 BP; 516 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Query Match 95.5%; Score 1464.2; DB 2; Length 1782;

Best Local Similarity 97.2%; Pred. No. 0;

Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Oy 1 ATGTTCAATCGAATTTTAAGGTAGAGAGTCTTAATGTAAGTACACCGAGCTGAGATT 60
Db 54 ATGTTCAATCGAATTTTAAGGTAGAGAGTCTTAATGTAAGTACACCGAGCTGAGATT 113
Oy 61 CAGTCCGTGTAACAATAAGCAACCAACCACTGTTTCAAGAAACAGAAATGGACCTAT 120
Db 114 CAGTCCGTGTAACAATAAGCAACCAACCACTGTTTCAAGAAACAGAAATGGACCTAT 173
Oy 121 CAGTGGATTGTCAAAACCCAAATCCGTCAACTACAAATTTAAACCAACCAATTTCCA 180
Db 174 CAGTGGATTGTCAAAACCCAAATCCGTCAACTACAAATTTAAACCAACCAATTTCTT 233
Oy 181 AAATGGGGGAGATGCTGTTGGGGTTGGGGTGGAAACCAACCGCTACCCGCTGAT 240
Db 234 AAATTAAGGGATATGCTGTTGGGGTGGGGTGGAAACCAACCGCTACCCGCTGAT 293
Oy 241 GTTATTGCTAACAGAGAGACATTTTATGAGCTACAAAGACCAAGATTCAACCAACCAAT 300
Db 294 GTTATTGCTAACAGAGAGACATTTTATGAGCTACAAAGACCAAGATTCAACCAACCAAT 353
Oy 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGGATCTTTCAGGAGAAAGAA 360
Db 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGGATCTTTCAGGAGAAAGAA 413
Oy 361 ATCTATGCCCATTCAGAGAGTGTGCTTCAATGTTAAATCCTGACGACATTTGTTGGG 420
Db 414 ATCTATGCCCATTCAGAGAGCTGCTTCAATGTTAAATCCTGACGACATTTGTTGGG 473
Oy 421 GGATGGATATTCAGCAACATGAACCTGGCTGATGTCATGCGCCAGGCAAAAGTGTGAAC 480
Db 474 GGATGGATATTCAGCAACATGAACCTGGCTGATGTCATGCGCCAGGCAAAAGTGTGAAC 533
Oy 481 ATCGATTTGACAGACAGTTGAGGCTTTTACATGATTCATGATTCATGCTCCCGGAATC 540
Db 534 ATCGATTTGACAGACAGTTGAGGCTTTTACATGATTCATGATTCATGCTCCCGGAATC 593
Oy 541 TACGACCCGGATTTTATGCTGCTCAACCAAGAGGAGCGGCCCAACAGTGTATTAAGGCG 600
Db 594 TATGACCCGGATTTTATGCTGCTCAACCAAGAGGAGCGGCCCAACAGTGTATTAAGGCG 653
Oy 601 ACAAGCAAGAGCAAGTTCAGCAATTCATCAAGACATCAAGGCTTTTAAGGAACCAACC 660
Db 654 ACAAGCAAGAGCAAGTTCAGCAATTCATCAAGACATCAAGGCTTTTAAGGAACCAACC 713
Oy 661 AAAGTGACAAAGGTGTTGCTGTTGACCTGTCACCTGCAACAGAGAGGTATTAAGCAATTTGTT 720
```

```

Db      714  AAAGTGACAAAGGTGGTGTGTAAGTGGAGCTGCCAACAAGAGAGTAAGTAATTTGGTT 773
Qy      721  GTAGGCTTAATATACCAATGAGAAATCTCTTGGCTGCTGTGACAGAAATAGGCTGAG 780
Db      774  GTGGGCTTAATATACCAATGAGAAATCTCTTGGCTGCTGTGACAGAAATAGGCTGAG 833
Qy      781  ATTTCTCTTCCACCTTGTATGCGATTCGCTGTGTATGGAATAATGTTCTTTCATTAT 840
Db      834  ATTTCTCTTCCACCTTGTATGCGATTCGCTGTGTATGGAATAATGTTCTTTCATTAT 893
Qy      841  GGAAGCCCTCAGAAACCTTTTGTACCAAGGCTGATGTATGATCTTTCACATCCGAGAAAC 900
Db      894  GGAAGCCCTCAGAAACCTTTTGTACCAAGGCTGATGTATGATCTTTCACATCCGAGAAAC 953
Qy      901  TTGATTTGGTGGAGATTAATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
Db      954  TTGATTTGGTGGAGATTAATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 1013
Qy      961  TTTCTTGTGGGGCTGTATCAAGCCATCTATAGTTAGTTACCAACCTTGGGAAAC 1020
Db      1014  TTCTCTTGTGGGGCTGTATCAAGCCATCTATAGTTAGTTACCAACCTTGGGAAAC 1073
Qy      1021  AATGATGTATGAATCTCTCGGCTCCACAAACCTTCGCTCCAAAGGAAATCTTCCAAAGC 1080
Db      1074  AATGATGTATGAATCTTTCGGCTCCACAAACCTTCGTTCCAAAGGAAATCTTCCAAAGC 1133
Qy      1081  AACGTTGTGAGATATGTGTCAACAGCAATGCCATCTCTATGAGCTGTGTAACTCCC 1140
Db      1134  AACGTTGTGAGATATGTGTCAACAGCAATGCCATCTCTATGAGCTGTGTAACTCCA 1193
Qy      1141  GACCATGTGTGTATTAATGATATGCTTACGTAAGGAGATGCAAGAGAGCATGAT 1200
Db      1194  GACCATGTGTGTATTAATGATATGCTTACGTAAGGAGATGCAAGAGAGCATGAT 1253
Qy      1201  GAGTACACTTCAGAGATATTCATGCGTGAAGAAACACCATTTGTTGCAACAACATGT 1260
Db      1254  GAGTACACTTCAGAGATATTCATGCGTGAAGAAACACCATTTGTTGCAACAACATGC 1313
Qy      1261  GAGGATTCCTCTTATGCTGCTCTATTAATCTTGAATGCTTGTCTTGTGAGCTGAGC 1320
Db      1314  GAGGATTCCTCTTATGCTGCTCTATTAATCTTGAATGCTTGTCTTGTGAGCTGAGC 1373
Qy      1321  ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACATCATTCACACCATGCT 1380
Db      1374  ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACATCATTCACACCATGCT 1433
Qy      1381  ACCATTCCTCAGCTATCTGACCAAGGCTCCTCTGTTCCACCGGATACACAGTGTGAAT 1440
Db      1434  ACCATTCCTCAGCTATCTGACCAAGGCTCCTCTGTTCCACCGGATACACAGTGTGAAT 1493
Qy      1441  GCATTGTCAAAGCAGCTGCAATGCTGAAACATATATGAGGCTTGTGTGATTTGCC 1500
Db      1494  GCATTGTCAAAGCAGCTGCAATGCTGAAACATATATGAGGCTTGTGTGATTTGCC 1553
Qy      1501  CCAGAGAAATPAATGATTTCTCGAGTACAGTGA 1533
Db      1554  CCAGAGAAATPAATGATTTCTCGAGTACAGTGA 1586

```

RESULT 12
AAV62443
ID AAV62443 standard; cDNA; 1533 BP.

AC AAV62443;
XX
XX 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
DE Soybean mutant myo-inositol 1-phosphate synthase cDNA.
XX
XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW phytic acid; ds.
XX

```

OS      Glycine max; line LR33.
XX
XX      MO9845448-A1.
XX
XX      15-OCT-1998.
XX
XX
XX      07-APR-1998; 98WO-US006822.
XX
XX      08-APR-1997; 97US-00835751.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX      Hitz WD, Sebastian SA;
XX
XX      WPI, 1998-568353/48.
XX
XX      P-PSDB; AAW79741.
XX
XX      Soybean plants containing altered myo-inositol-1-phosphate gene - useful
XX      for generating plants with altered levels of e.g. raffinose, stachyose,
XX      phytic acid, etc.
XX
XX      Example 5; Page 48-49; 63pp; English.
XX
XX      This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
XX      inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
XX      clone was isolated from a cDNA library of soybean line LR33 by PCR
XX      amplification (see AAV62441-42). Line LR33 was obtained by chemical
XX      mutagenesis of wild-type soybean genome and as a reduced raffinose
XX      saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
XX      acid, raffinose and stachyose. Sequencing revealed a single base change
XX      mutation (G to T at base 1241) in the LR33 sequence when compared to the
XX      wild-type sequence (see AAV62440). The mutation results in a seed
XX      phenotype of very low raffinose saccharide sugars, very high sucrose and
XX      low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
XX      raffinose saccharide, sucrose, phytic acid and inorganic phosphate
XX      content of soybean seeds, leading to useful soybean products, e.g. a seed
XX      phytic acid content of less than 17 ug/g, a seed content of raffinose and
XX      stachyose combined of less than 14.5 ug/g, and a seed sucrose content
XX      greater than 200 ug/g. (Updated on 17-Oct-2003 to standardise OS field)
XX
XX      Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
XX
XX      Query Match      95.4%; Score 1462.6; DB 2; Length 1533;
XX      Best Local Similarity 97.1%; Pred. No. 0;
XX      Matches 1489; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

```

Qy      1  ATGTTCAATGAGAAATTTTAAAGTGAAGAGTCTTAATGTGAATACCGAGACTGAGATT 60
Db      1  ATGTTCAATGAGAAATTTTAAAGTGAAGAGTCTTAATGTGAATACCGAGACTGAGATT 60
Qy      61  CAGTCGCTGTACAACTACGAAACCAACGAACTTGTTCAGAGAAACAGAAATGCACTTAT 120
Db      61  CAGTCGCTGTACAACTACGAAACCAACGAACTTGTTCAGAGAAACAGAAATGCACTTAT 120
Qy      121  CAGTGGATTTGTAAACCCAAATCCGTCACACTCAATTTTAAACCAACCAACGATTTCCA 180
Db      121  CAGTGGATTTGTAAACCCAAATCCGTCACACTCAATTTTAAACCAACCAACGATTTCTT 180
Qy      181  AAATTGGGGGTGATGCTTGTGGGTGGAGTGAACCAACGAGCTTACCTCAACCGGTGAT 240
Db      181  AAATTGGGGGTGATGCTTGTGGGTGGAGTGAACCAACGAGCTTACCTCAACCGGTGAT 240
Qy      241  GTTATTTGCTAACAGAGAGACATTTTCAATGGGCTACAAAGAGACAAATTCACAGCCAT 300
Db      241  GTTATTTGCTAACAGAGAGACATTTTCAATGGGCTACAAAGAGACAAATTCACAGCCAT 300
Qy      301  TACTTTGGCTCCCTCAACCCAGAGCTCAGACTATTCAGAGTGTGATCTTTCAGAGGAGAA 360
Db      301  TACTTTGGCTCCCTCAACCCAGAGCTCAGACTATTCAGAGTGTGATCTTTCAGAGGAGAA 360
Qy      361  ATCTATGCCCATTCAGAAAGTCTGCTTCCAAATGTTAATCTCGACGACATTTGTGGG 420
Db      361  ATCTATGCCCATTCAGAAAGTCTGCTTCCAAATGTTAATCTCGACGACATTTGTGGG 420

```

421 GGATGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGSCAAAGGTGTTGAC 480
 421 GGATGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGSCAAAGGTGTTGAC 480
 481 ATCGATTTGACAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC 540
 481 ATCGATTTGACAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC 540
 541 TACGACCCGGATTTGATGCTGSCCAACCAAGAGGCGGCCAACAAAGTGATTAAGGAC 600
 541 TATGACCCGGATTTGATGCTGSCCAACCAAGAGGCGGCCAACAAAGTGATTAAGGAC 600
 601 ACAAAAGCAAGACCAAGTTCCAGCAAAATCATCAAAAGCATCAAGACGCTTTAAGGAACCA 660
 601 ACAAAAGCAAGACCAAGTTCAACAAATCATCAAAAGCATCAAGACGCTTTAAGGAACCA 660
 661 AAAGTGACAGAGGTGTTCTCTGAGACTGCCAACAGAGAGTATGCAATTTGGTT 720
 661 AAAGTGACAGAGGTGTTCTCTGAGACTGCCAACAGAGAGTATGCAATTTGGTT 720
 721 GTAGGCTTATGACACCATGAGAAATCTCTGGCTGCTGAGAGAGAAATGAGGCTGAG 780
 721 GTAGGCTTATGACACCATGAGAAATCTCTGGCTGCTGAGAGAGAAATGAGGCTGAG 780
 781 ATTTCTCTTCCACCTTGATGCCATGCTGCTGATGAGAAATGTTCTTTCAATTAAT 840
 781 ATTTCTCTTCCACCTTGATGCCATGCTGCTGATGAGAAATGTTCTTTCAATTAAT 840
 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTCCATCCGAGAACACT 900
 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTCCATCCGAGAACACT 900
 901 TTGATTTGGTGAAGTACTTCAAGAGTGGTCAAGACCAAAATGAAATCTGTGTTGGTAT 960
 901 TTGATTTGGTGAAGTACTTCAAGAGTGGTCAAGACCAAAATGAAATCTGTGTTGGTAT 960
 961 TTTCTTGGGGGCTGATATCAAGCCACATCTATAGTTAGTTACCAACCATCTGGGAAAC 1020
 961 TTTCTTGGGGGCTGATATCAAGCCACATCTATAGTTAGTTAGTTACCAACCATCTGGGAAAC 1020
 1021 AATGATGATGATCTCTGGCTCCACAAACCTTCCGCTCCAGAGAAATCTTCCAGAGAC 1080
 1021 AATGATGATGATCTCTGGCTCCACAAACCTTCCGCTCCAGAGAAATCTTCCAGAGAC 1080
 1081 AAGTGTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 1081 AAGTGTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 1141 GACCATGTTGTTATTAAGTATGCTTACGTAAGGAGATGACAGAGACCAATGAT 1200
 1141 GACCATGTTGTTATTAAGTATGCTTACGTAAGGAGATGACAGAGACCAATGAT 1200
 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACACATTTGTTTGCACACATAT 1260
 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACACATTTGTTTGCACACATAT 1260
 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACACATTTGTTTGCACACATAT 1260
 1261 GAGGATTCCTTTAGACCTGCTCTATTAATCTTGGAATGATGATGATGATGATGATGATGATGAT 1320
 1261 GAGGATTCCTTTAGACCTGCTCTATTAATCTTGGAATGATGATGATGATGATGATGATGATGAT 1320
 1321 ACTAGATTCAGATTTAAAGCTGAAATGAGGAAATTCACATCTTCCAGCCAGTTGCT 1380
 1321 ACTAGATTCAGATTTAAAGCTGAAATGAGGAAATTCACATCTTCCAGCCAGTTGCT 1380
 1381 ACCATTTCTAGCTATTTGACCAAGGCTCTCTGTTTCCACCGGATACACCAAGTGGAT 1440
 1381 ACCATTTCTAGCTATTTGACCAAGGCTCTCTGTTTCCACCGGATACACCAAGTGGAT 1440
 1441 GGATTTGCAAGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCT 1500
 1441 GGATTTGCAAGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCT 1500

1501 CCAGAGATTAACATGATTTCTCGAGTACAGTGA 1533
 1501 CCAGAGATTAACATGATTTCTCGAGTACAGTGA 1533
 RESULT 13
 AD014494 standard; cDNA; 1533 BP.
 AD014494;
 23-SEP-2004 (first entry)
 Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.
 Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 inorganic phosphate; mutant.
 Glycine max.
 Synthetic.
 Key Location/Qualifiers
 CDS 1..1533
 /tag= a
 /product= "Mutant soybean myo-inositol 1-phosphate
 synthase #1"
 mutation replace(1241,G)
 /tag= b
 US2004128713-A1.
 01-JUL-2004.
 21-NOV-2003; 2003US-00718952.
 08-APR-1997; 97US-00835751.
 07-APR-1998; 98WO-US006822.
 26-APR-1999; 99US-00299315.
 11-MAR-2002; 2002US-00025003.
 (HITZ/) HITZ W D.
 (SEBA/) SEBASTIAN S A.
 (GRAC/) GRACE D J.
 (STRE/) STREIT L G.
 Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 WPI; 2004-53135/51.
 P-PSDB; AD014495.
 New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 useful for producing plants with decreased raffinose, stachyose, and
 phytic acid and increased sucrose, leading to valuable and useful soybean
 products.
 Claim 10; SEQ ID NO 5; 48bp; English.
 The invention relates to a nucleic acid fragment encoding a soybean myo-
 inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 synthase having a decreased capacity for the synthesis of myo-inositol 1-
 phosphate. The invention also relates to a chimeric gene operably linked
 to suitable regulatory sequences, where expression of the chimeric gene
 results in a decrease in expression of an endogenous or native gene
 encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 comprising the chimeric gene, a method of making the soybean plant, a
 seed of the soybean plant, a soy protein product derived from the
 processing of soybean seeds, a method of making or producing a soy
 protein product and a method of using a soybean plant homozygous for at
 least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 nucleic acid fragment and methods are useful for producing plants with
 decreased raffinose, stachyose and phytic acid content and increased

CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 95.4%; Score 1462.6; DB 12; Length 1533;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1489; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

QY 1 ATGTCATCGAGAAATTTAAAGTAGAGAGCTCTAATGTGAAGTACACCGAGCTGAGATT 60
DB 1 ATGTCATCGAGAAATTTAAAGTAGAGAGCTCTAATGTGAAGTACACCGAGCTGAGATT 60
QY 61 CAGTCGCTGTAACAACCAACCACTTGTTCAGAGAAACAGAAATGGCACTTAT 120
DB 61 CAGTCGCTGTAACAACCAACCACTTGTTCAGAGAAACAGAAATGGCACTTAT 120
QY 121 CAGTGAATGTCAAAACCAACCACTGTAACCAATTTAAACCAACCACTGTTCCA 180
DB 121 CAGTGAATGTCAAAACCAACCACTGTAACCAATTTAAACCAACCACTGTTCC 180
QY 181 AAATTGGGGGTGATGCTTGGGGTGGGGTGAACAACGGCTCTACCCCTCACCGGTGGT 240
DB 181 AAATTGGGGGTGATGCTTGGGGTGGGGTGAACAACGGCTCTACCCCTCACCGGTGGT 240
QY 241 GTTATTGCTTAACAGAGACATTTTATGAGCTTACAAAGACAAAGATTCAACAGCCAA 300
DB 241 GTTATTGCTTAACAGAGACATTTTATGAGCTTACAAAGACAAAGATTCAACAGCCAA 300
QY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGGAATCCCTTCAAGGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGGAATCCCTTCAAGGGAGAA 360
QY 361 ATCTATGCCCATTTCAAGAGCTGCTTCAATGCTTAACTCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGCTGCTTCAATGCTTAACTCTGACGACATTTGTTGGG 420
QY 421 GGATGGGAATATCAGCAACATGAACCTGCTGATGCAATGGCCAGGGCAAAAGTGTGAC 480
DB 421 GGATGGGAATATCAGCAACATGAACCTGCTGATGCAATGGCCAGGGCAAAAGTGTGAC 480
QY 481 ATCGATTTCAGAAAGAGTTGAGGCTTTCATGGAATTCATGCTCCCGGGAATC 540
DB 481 ATCGATTTCAGAAAGAGTTGAGGCTTTCATGGAATTCATGCTCCCGGGAATC 540
QY 541 TACGACCCCGGATTTCTATGCTGCAACCAAGAGACGCTGCCAACACGTCATTAAGGAC 600
DB 541 TACGACCCCGGATTTCTATGCTGCAACCAAGAGACGCTGCCAACACGTCATTAAGGAC 600
QY 601 ACAAAAGCAAGACAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAGCAACC 660
DB 601 ACAAAAGCAAGACAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAGCAACC 660
QY 661 AAAAGTGAACAAGGTGTTCCGTTGAGTGCACACAGAGAGTATGCAATTTGGT 720
DB 661 AAAAGTGAACAAGGTGTTCCGTTGAGTGCACACAGAGAGTATGCAATTTGGT 720
QY 721 GTAGGCTTAAATGACACATGAGAAATCTTTGCTGCTGTGAGACAGAAATGAGCTGAG 780
DB 721 GTAGGCTTAAATGACACATGAGAAATCTTTGCTGCTGTGAGACAGAAATGAGCTGAG 780
QY 781 ATTTCTCTTCACTTGTATGCAATGCTGCTGTGATGAAATGTTCTTTTCAATTAAT 840
DB 781 ATTTCTCTTCACTTGTATGCAATGCTGCTGTGATGAAATGTTCTTTTCAATTAAT 840
QY 841 GGAAGCCCTCAGAAACCTTTGTAACAGGGCTGATGATCTTGCCATCCGCGAGAACT 900
DB 841 GGAAGCCCTCAGAAACCTTTGTAACAGGGCTGATGATCTTGCCATCCGCGAGAACT 900
QY 901 TTGATTGCTGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 960
DB 901 TTGATTGCTGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGAT 960

```

```

QY 961 TTCTCTGGGGGGCTGCTGATCAAGCCAACTATATGATTAGTACAAACATCTGGGAAC 1020
DB 961 TTCTCTGGGGGGCTGCTGATCAAGCCAACTATATGATTAGTACAAACATCTGGGAAC 1020
QY 1021 AATGATGTATGATATCTCGGCTCCCAAAACCTTCGCTCCAAAGAAATCTTCAAGAC 1080
DB 1021 AATGATGTATGATATCTTCGGCTCCCAAAACCTTCGCTCCAAAGAAATCTTCAAGAC 1080
QY 1081 AACGTTGTTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGCTGTAACATCCC 1140
DB 1081 AACGTTGTTGATATGCTCAACAGCAATGCCATCTCTATGAGCTGCTGTAACATCCA 1140
QY 1141 GACATGTTGTTGTTATTAAGTATGCTTACGTAAGGGGATAGCAAGAGCCATGAT 1200
DB 1141 GACATGTTGTTGTTATTAAGTATGCTTACGTAAGGGGATAGCAATGAGCCATGAT 1200
QY 1201 GAGTACATTCAGAGATATTCATGAGGTGAAGAAACACCATTTGTTGCAACACATGT 1260
DB 1201 GAGTACATTCAGAGATATTCATGAGGTGAAGAAACACCATTTGTTGCAACACATGC 1260
QY 1261 GAGGATTCCTTTAGCTGCTCTTATCTTGAATTTGACTTGTCTTCTGAGCTGAC 1320
DB 1261 GAGGATTCCTCTTAGCTGCTCTTATCTTGAATTTGACTTGTCTTCTGAGCTGAC 1320
QY 1321 ACTAGATCACTTAAAGCTGAATAAGGAAATTCACATCTATTCACACCAAGTTGCT 1380
DB 1321 ACTAGATCACTTAAAGCTGAATAAGGAAATTCACATCTATTCACACCAAGTTGCT 1380
QY 1381 ACCATTCAGCTATCTGACCAAGGCTCTGTTTCCACCGGGTACACAGTGTGAT 1440
DB 1381 ACCATTCAGCTATCTGACCAAGGCTCTGTTTCCACCGGGTACACAGTGTGAT 1440
QY 1441 GCATTGTCAAACAGCTGCAATGCTGAAACCATTAATGAGGCTTGTGATTTGGC 1500
DB 1441 GCATTGTCAAACAGCTGCAATGCTGAAACCATTAATGAGGCTTGTGATTTGGC 1500
QY 1501 CCAGAGATTAATGATTTCTGATTAACAGTGA 1533
DB 1501 CCAGAGATTAATGATTTCTGATTAACAGTGA 1533

RESULT 14
ADS8197
ID ADS8197 standard; cDNA; 1533 BP.
XX
AC ADS8197;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
OS Glycine max; line LR33.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..1533
FT CDS /product= "myo-inositol 1-phosphate synthase"
FT mutation /tag= b
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1996; 98WO-US006822.

```


DB 1381 ACCATCTCAGTACCTCAACCAAGGCTCTCTGTTCCACCGGGTACACGAGTGTGAAT 1440
QY 1441 GCATTGCAAGAGGCTGCATGTCGAAACATATATGAGGGCTGTGTGATGGCC 1500
DB 1441 GCATTGCAAGAGGCTGCATGTCGAAACATATATGAGGGCTGTGTGATGGCC 1500
QY 1501 CCAGAGATTAACATGATTCGTGAGTACAAGTGA 1533
DB 1501 CCAGAGATTAACATGATTCGTGAGTACAAGTGA 1533
RESULT 15
AA90402
ID AAX90402 standard; cDNA to mRNA, 1950 BP.
XX AAX90402;
XX
XX 24-SEP-1999 (first entry)
XX
XX Nicotiana paniculata INPS encoding cDNA.
XX
XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
XX water stress; resistance; ds.
XX
XX Nicotiana paniculata.
XX
XX JP1187879-A.
XX
XX 13-JUL-1999.
XX
XX 26-DEC-1997; 97JP-00359773.
XX
XX 26-DEC-1997; 97JP-00359773.
XX
XX (N1SB) JAPAN TOBACCO INC.
XX
XX WPI; 1999-451546/38.
XX P-PSDB; AAY24477.
XX
XX New INPS gene derived from Nicotiana genus plant - useful for conferring
XX resistance to water stress to plants.
XX
XX
XX Claim 2; Page 6-8; Bpp; Japanese.
XX
XX The present sequence encodes Nicotiana paniculata inositol monophosphate
XX synthase (INPS), designated NpINPS1. INPS can be used to confer water
XX stress resistance to a plant
XX
XX Sequence 1950 BP; 570 A; 377 C; 425 G; 578 T; 0 U; 0 Other:
SQ
Query Match 71.6%; Score 1097.8; DB 2; Length 1950;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 1261; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
QY 1 AGGTCATCGAGAAATTTAAGGTAGAGAGTCCTAATGTGAAGTACCGAGCTGAGATT 60
DB 92 AGGTTATATGAGAAATTTAAGGTAGAGAGTCCTAATGTGAAGTACCGAGTGAAT 151
QY 61 CAGTCGTGTACAACTACGAAACCAACGAACTGTTCAGAGAAACGAAATGCACTAT 120
DB 152 CACTGTGTATGATATCAAAACCACTGATGTTATCATGATGAGAAAAATGGACATAT 211
QY 121 CAGTGTATGTCAAAACCAATCCGTCATCTACCAATTTAAACCAACCACTGTTCA 180
DB 212 CAGTGCACGTCAAGCTAAGACTGTCAAAATATGATTCAGACTGATGTTATGCC 271
QY 181 AATTTGGGGGTATGCTTGTGGGTGGGAGAAACAAGGCTCTACCCCTCAACCGGTGAT 240
DB 272 AATTTGGGGGTATGCTTGTGGGTGGGAGAAACAATGTTCAACCTTGACCGGTGAT 331
QY 241 GTTATTTGCTAACAGAGACATTTTCATGAGGCTACAAAGACAAAGATTCAAGCCAT 300
DB 332 GTTATTTGCTAACAGAGAAAGAAATTTTCATGAGGCTCAACAAAGATTAAGTGCACAAGCCAT 391

QY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTCCAGTTGATTCCTTCCAGGAGAGAA 360
DB 392 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTCCAGTTGATTCCTTCCAGGAGAG 451
QY 361 ATCTATGCCCAATTCAGAGTCTGCTTCCATGTTATCTGACGACATGTTGTTGG 420
DB 452 ATCTATGCTCCATTTAAAGCTCTCTTCCATGTTATCTGACGATACGATGTTGGA 511
QY 421 GGAATGGATATCAAGCAATGAACTGAGTATGCCATGAGGCAAGGCAAGGTTTGAC 480
DB 512 GGAATGGATATCAAGCAATGAACTGAGTATGCCATGAGGCAAGGCAAGGTTTGAC 571
QY 481 ATGATTTGCAAGAGAGTGGAGGCTTATCATGAAATCATGTTGTTCACTCCCGCAATC 540
DB 572 ATGATTTGCAAGAGAGTGGAGGCTTATCATGAAATCATGTTGTTCACTCCCGCAATC 631
QY 541 TAGACCCGGAATTTCTATGCTGCAACCAAGAGAGGCTGCCCAACAGTATTAAGGCG 600
DB 632 TAGACCCGGAATTTCTATGCTGCTATCAACAGGAGTCACTGCCCAACAGTATTAAGG 691
QY 601 ACAAGCAAGCAAGTTCAGCAAAATCATCAAGACATCAAGGCTTTAAGGAGCCACC 660
DB 692 ACCAAGAAAGCAAAATGATCAAAATCATTAAGATATTAAGGAGTTTAAGGAAAGAAC 751
QY 661 AAGTGGACAGAGTGTGCTCTGTCGATGCCCAACAGAGATATGCAATTTGTT 720
DB 752 AAGTGGACAGAGTGTGATTTGAGACTGTCAACAGTAAAGATACGATATGTT 811
QY 721 GTAGGCTTAAATGACACATGAGAGATCTTGGGCTGCTGTGAGCAAGAAATGAGCTGAG 780
DB 812 GTTGGACCTTAATGACATGAGAGAACTCTTGGCTCTGTGAGCAAGAAATGAGCTGAA 871
QY 781 ATTTCTCTTCCACCTTGATGCAATGCTGTCGATGAGAAATGTTCTTTCAATTAAT 840
DB 872 ATATCTCTTCCACCTTGATGCTATGCTGTCGATGCTGATTTGAATGCTTTTATTAAT 931
QY 841 GGAAGCCCTCAAGAACTTTTATGAGGCTGATTTGATCTTCCATGCCAGAACT 900
DB 932 GGAAGCCCTCAAGAACTTTTATGAGGCTGATTTGATCTTCCATGCCAGAACT 991
QY 901 TTGATTTGATGAGATGATCTTCAAGAGTGTGAGCAACCAAAATGTAATCTGTTGTTGAT 960
DB 992 TTGATTTGATGAGATGATCTTCAAGAGTGTGAGCAACCAAAATGTAATCTGTTGTTGAT 1051
QY 961 TTTCTTGGGGGCTGTATCAAGCAACATCTATAGTTAGTTACCAACATCTGGGAAAC 1020
DB 1052 TTTCTTGGGGGCTGTATCAAGCAACATCTATAGTTAGTTACCAACATCTGGGAAAC 1111
QY 1021 AATGATGATGATATCTCTGAGCTCAACAACTTCCGCTCAAGGAAATTTCCAGAGC 1080
DB 1112 AATGATGATGATATCTCTGAGCTCAACAACTTCCGCTCAAGGAAATTTCCAGAGC 1171
QY 1081 AAGCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1172 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
QY 1141 GACATGTTGTTGATTTAAGTATGCTTACGTAAGGAGATGAGAGACATGAT 1200
DB 1232 GACATGTTGTTGATTTAAGTATGCTTACGTAAGGAGATGAGAGACATGAT 1291
QY 1201 GAGTACATTCAGAGATATTCATGAGGAGAAAGAACACCATTTGTTGCAACACATGT 1260
DB 1292 GAGTACATTCAGAGATATTCATGAGGAGAAAGAACACCATTTGTTGCAACACATGT 1351
QY 1261 GAGATTCCTTTTATGCTGCTCTATTTATCTTGAATGCTTCTTCTGAGCTGAGC 1320
DB 1352 GAGATTCCTTTTATGCTGCTCTATTTATCTTGAATGCTTCTTCTGAGCTGAGC 1411
QY 1321 ACTAGAAATCAAGTTTAAAGCTGAAATATGAGGAAATTTCACTATTCACACCACTTCT 1380
DB 1412 ACCGCAATTCAGCTCAAGGCTGAAAGAGAGGATTAAGTTCACTCTTCAACCCGCTGCT 1471

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 20:51:42 ; Search time 6686.92 Seconds
(without alignments)
11108.537 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgtcaccggaattttaa.....tgattctcgagtaacaagrga 1533

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmb1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1464.2	95.5	1739	8	AY038802 Glycine m
2	1464.2	95.5	1782	8	BD075266 Soybean p
3	1462.6	95.4	1533	6	BD075269 Soybean p
4	1456.2	95.0	1791	8	AF293970 Glycine m
5	1097.8	71.6	1931	8	AB032073 Nicotiana
6	1097.8	71.6	1950	6	E27176 Novel INPS
7	1093	71.3	1959	8	AB009881 Nicotiana
8	1089	69.7	1954	8	BT013759 Lycopersi
9	1061	68.2	1959	8	BT013505 Lycopersi
10	1045	68.2	1845	8	AF284065 Sesamum i
11	1044.6	68.1	1978	8	CPINOLG C.paradis
12	1027.4	67.0	2053	8	U32531 Meembryant
13	1020.4	66.6	1986	8	AF433879 Suaeda ma
14	1016.2	66.3	1845	8	AY028259 Avicennia
15	1000.2	65.2	1781	6	AX054630 Sequence
16	1000.2	65.2	1781	6	BN06307 Brassaica na
17	998.6	65.1	1533	6	CO805008 Sequence
18	998.6	65.1	1533	6	AX506743 Sequence
19	998.6	65.1	1533	8	AY143904 Arabidops

20	998.6	65.1	1863	8	AY054202 Arabidops
21	998.6	65.1	1864	8	AY053415 Arabidops
22	997	65.0	1538	8	AF120146 Triticum
23	997	65.0	1538	8	AF120147 Triticum
24	997	65.0	1538	8	AF120148 Triticum
25	990.6	64.6	1564	8	AY096554 Arabidops
26	990.6	64.6	1863	8	AY065415 Arabidops
27	989	64.5	1890	8	ATU30250 Arabidops
28	987.2	64.4	1870	8	AF372954 Arabidops
29	985.8	64.3	1535	6	BD073472 Regulated
30	985.8	64.3	1665	6	BD073470 Regulated
31	985.8	64.3	1665	8	AF056326 Zea mays
32	985.6	64.3	1567	8	BT001931 Arabidops
33	985.6	64.3	1837	8	AY085931 Arabidops
34	984.2	64.2	1533	8	AY323824 Xerophyta
35	984.2	64.2	1959	6	AX151207 Sequence
36	982.4	64.0	1921	8	ATU04876 Arabidops
37	981	64.0	1914	8	AK103501 Arabidops
38	981	64.0	1915	8	AK058750 Oryza sat
39	979.4	63.9	1868	8	AB012107 Oryza sat
40	979.4	63.9	1931	6	AR137808 Sequence
41	979.4	63.9	1931	6	AR170167 Sequence
42	976	63.7	1871	8	PVU38920 Phaseolus v
43	974.6	63.6	1936	8	AB059557 Avena sat
44	973.2	63.5	1930	8	AY035045 Arabidops
45	969.8	63.3	2152	8	AF056325 Hordeum v

ALIGNMENTS

RESULT 1
LOCUS AY038802 1739 bp mRNA linear PLN 16-JUL-2001
DEFINITION Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.
ACCESSION AY038802
VERSION AY038802.1 GI:14764465

KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)

Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
1 (bases 1 to 1739)

AUTHORS Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.

TITLE Biochemical and Molecular Characterization of a Mutation that

Confers a Decreased Raffinose and Phytic Acid Phenotype on

Soybean Seeds

Unpublished

2 (bases 1 to 1739)

AUTHORS Carlson,T.J. and Hitz,W.D.

TITLE Direct Submission

Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402,

Wilmington, DE 19880-0402, USA

Location/Qualifiers

1..1739

/organism="Glycine max"

/mol_type="mRNA"

/culivar="Wye"

/db_xref="taxon:3847"

/dev_stage="developing embryo"

11..1543

/EC_number="5.5.1.4"

/function="cyclizes glucose 6-phosphate to

L-myo-inositol-1-phosphate"

/note="m1PS-1A"

/codon_start=1

/product="myo-inositol-1-phosphate synthase"

/protein_id="AAK72098.1"

/db_xref="GI:14764466"

/translation="MEIENFVCEGCVNYTETIOSVYVYETTELVHNRNGTYOMIV
KPKSVKTEFKTNHVPKLGVMVGNGNGSTLTGVIANREGISMATKDKIQANVF

GSUOASAIRVSGFOGEBIYAFKSLBPVNPDDIVFGMDISNNKLNADAMARAFVD
IDLOKOLPYMESMLPLPIGYDPRDIANNOBERANNVYKGTQBOVOOIIIDIXAFKE
ATKDYRUVVLTAMTERYSNLVYGVNDPMENLILAVDNBAEISSTYXIAACHVENV
PFINSPOVTEPGLIDILAIANTLIGDDENFSGOTKXSVLVDLVAGAKPISIVS
YNHLKNDGOMNLAPQOTFRSKISXSNVDPMVSNALYEBGEHPDHVAVIKVYVY
GDSKAMBEYSEIEMGKSTIVLANTCEDSLAIIIDIVLAEISLSTREKENE
KRFHSFHPVATILSYLTAKAPlVPBGTPVNNALSKORAMLENIMRACVGLAPENNMLE
YK"

ORIGIN

Query Match 95.5%; Score 1464.2; DB 8; Length 1739;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 ATGTTTCATCGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAATACACCGAGCTGAGATT 60
DB 11 AGTTTCATCGAGAAATTTTAAAGTAGAGTCTTAATGTGAATACACCGAGCTGAGATT 70
QY 61 CAGTCGGGTACACGAGAAACACCGAACTTTGTTACGAGAACGGAATGGCACTTAT 120
DB 71 CAGTCGGGTACACGAGAAACACCGAACTTTGTTACGAGAACGGAATGGCACTTAT 130
QY 121 CAGTGAATGTCAAACCCAAATCTGTCAATACGAATTTAAACCAACATCATGTCTCT 180
DB 131 CAGTGAATGTCAAACCCAAATCTGTCAATACGAATTTAAACCAACATCATGTCTCT 190
QY 181 AAATTTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTCACCGTGT 240
DB 191 AAATTTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTCACCGTGT 250
QY 241 GTTATTGTCTAACAGAGACATTTTCATGGGCTAACAAAGACAGATTCAACAAACCAAT 300
DB 251 GTTATTGTCTAACAGAGAGGACATTTTCATGGGCTAACAAAGACAGATTCAACAAACCAAT 310
QY 301 TACTTTGGCTCCCTCAACCCAGCCTCAAGTATTCAGATTGATCTTCCAGGAGAGAA 360
DB 311 TACTTTGGCTCCCTCAACCCAGCCTCAAGTATTCAGATTGATCTTCCAGGAGAGAA 370
QY 361 ATCTATGCCCCATTCAAGAGTCTGCTTCCAATGGTTAACTCTGACGACATTTGTGTTGG 420
DB 371 ATCTATGCCCCATTCAAGAGCCTGCTTCCAATGGTTAACTCTGACGACATTTGTGTTGG 430
QY 421 GGATGGGATATACGAAACATGAACCTGCTGATGTCATGGCCAGGCAAAAGGTGTTTAC 480
DB 431 GGATGGGATATACGAAACATGAACCTGCTGATGTCATGGCCAGGCAAAAGGTGTTTAC 490
QY 481 ATCGATTGTCAGAAAGCAGTTGAGGCTTTACATGAATTCATGATGCTTCCCGGAATC 540
DB 491 ATCGATTGTCAGAAAGCAGTTGAGGCTTTACATGAATTCATGATGCTTCCCGGAATC 550
QY 541 TACGACCCCGAATTTCTTGTCTGCCAACAAAGAGAGGCTGCCAACACGTGATTTAAGGC 600
DB 551 TACGACCCCGAATTTCTTGTCTGCCAACAAAGAGAGGCTGCCAACACGTGATTTAAGGC 610
QY 601 ACAAAGCAAGACAGATTGAGCAAAATCATCAAAAGACATCAAGCCGTTTAAAGAAAGCAC 660
DB 611 ACAAAGCAAGACAGATTGAGCAAAATCATCAAAAGACATCAAGCCGTTTAAAGAAAGCAC 670
QY 661 AAAGTGACAAAGGTGTTCTCTGTGAGTCCCAACAGAGAGGTATAGCAATTTGGTT 720
DB 671 AAAGTGACAAAGGTGTTCTCTGTGAGTCCCAACAGAGAGGTATAGCAATTTGGTT 730
QY 721 GTAGGCTTAATGACACCATGAGATCTCTTGGCTGCTGTGACAGAAATAGAGGTGAG 780
DB 731 GTAGGCTTAATGACACCATGAGATCTCTTGGCTGCTGTGACAGAAATAGAGGTGAG 790
QY 781 ATTTCTCTTCCACCTTGATGCAATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
DB 791 ATTTCTCTTCCACCTTGATGCAATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 850
QY 841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCCATTCGAGAAACACT 900
DB 851 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCCATTCGAGAAACACT 910

QY 901 TTGATTGGTGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGGTTGAT 960
DB 911 TTGATTGGTGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGGTTGAT 970
QY 961 TTTCTTGTGGGGGCTGTGATGACCAACATCTATAGTTAGTTAACAACCATCTGGGAAC 1020
DB 971 TTTCTTGTGGGGGCTGTGATGACCAACATCTATAGTTAGTTAACAACCATCTGGGAAC 1030
QY 1021 AATGATGATGATCTCTCGGCTTCACAAACCTTCCGCTCCAGGAATCTCCAAAGAC 1080
DB 1031 AATGATGATGATCTCTCGGCTTCACAAACCTTCCGCTCCAGGAATCTCCAAAGAC 1090
QY 1081 AACGTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1091 AACGTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1150
QY 1141 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1151 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
QY 1201 GAGTACATTTGAGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1211 GAGTACATTTGAGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1270
QY 1261 GAGATTCCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1271 GAGATTCCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330
QY 1321 ACTAGAAATCAAGTTTAAAGCTGAAATGAGGAAATTTCACTCATTCACCCAGTTGCT 1380
DB 1331 ACTAGAAATCAAGTTTAAAGCTGAAATGAGGAAATTTCACTCATTCACCCAGTTGCT 1390
QY 1381 ACCATTCTCAGCTATCTACCAAGGCTCTCTGTTTCCACCGGGTACACAGTGGTGAAT 1440
DB 1391 ACCATTCTCAGCTATCTACCAAGGCTCTCTGTTTCCACCGGGTACACAGTGGTGAAT 1450
QY 1441 GCATTGTCAAAGCAGCAGTGCATGCTGGAATAACATATGAGGCTTGTGTTGATTTGCC 1500
DB 1451 GCATTGTCAAAGCAGCAGTGCATGCTGGAATAACATATGAGGCTTGTGTTGATTTGCC 1510
QY 1501 CCAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533
DB 1511 CCAGAAATTAACATGATTTCTGAGTACAAAGTGA 1543

RESULT 2
BD075266 1782 bp DNA linear PAT 27-AUG-2002
LOCUS BD075266
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION BD075266
VERSION BD075266.1 GI:22620869
KEYWORDS JP 2001519665-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1782)
AUTHORS Hitz,W.D. and Sebastian,S.A.
TITLES Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid
JOURNAL Patent: JP 2001519665-A 1 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR13
PN JP 2001519665-A/1
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC
Strandness: Double;
CC Topology: linear;
CC Soybean plant producing seeds with reduced

Levels of raffinose
CC and phytic acid Location/Qualifiers
FH Key 54.1586.
FT CDS Location/Qualifiers
1.1782
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 95.5%; Score 1464.2; DB 6; Length 1782;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

1 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
54 ATGTTTCATGAGAAATTTTAAAGTAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 113
61 CAGTCCGCTGTAACAACCAACCAACCACTTGTTCAGAGAAACAGAAATGGCACTTAT 120
114 CAGTCCGCTGTAACAACCAACCAACCACTTGTTCAGAGAAACAGAAATGGCACTTAT 173
121 CAGTGGATGTCAAAACCAATCCGCTCACTACCAATTTTAAACCAACCAACCACTTAT 180
174 CAGTGGATGTCAAAACCAATCCGCTCACTACCAATTTTAAACCAACCAACCACTTAT 233
181 AAATTTGGGGGTATGCTTGTGGGTGGGGTGAACCAACGAGCTTACCTCAACCGGTGCT 240
224 AAATTTAGGGGTATGCTTGTGGGTGGGGTGAACCAACGAGCTTACCTCAACCGGTGCT 293
241 GTTATTTGCTAACAGAGAGACATTTTCAATGGGCTTCAAGAGAACAGATTCAAGCAAT 300
294 GTTATTTGCTAACAGAGAGACATTTTCAATGGGCTTCAAGAGAACAGATTCAAGCAAT 353
301 TACTTTGGCTCCCTCAACCAACCACTCACTATTCAGTTTGAATCTTCCAGGAGAGGAA 360
354 TACTTTGGCTCCCTCAACCAACCACTCACTATTCAGTTTGAATCTTCCAGGAGAGGAA 413
361 ATCTATGCCCCATTGAGAGAGTGTCTTCCATGTGTTAATCTGACGACATTTGTTGGG 420
414 ATCTATGCCCCATTGAGAGAGTGTCTTCCATGTGTTAATCTGACGACATTTGTTGGG 473
421 GGATGGGATATACGACATGAACTGTGCTGATGCCATGCGCCAGGCAAAAGGTGTTGAC 480
474 GGATGGGATATACGACATGAACTGTGCTGATGCCATGCGCCAGGCAAAAGGTGTTGAC 533
481 ATGATTTGGACAGACAGTTGAGGCTTATCATGGAATCCATGGTTCCACTCCCGGAAATC 540
534 ATGATTTGGACAGACAGTTGAGGCTTATCATGGAATCCATGGTTCCACTCCCGGAAATC 593
541 TACGACCCGAGATTTGATGCTGCTCCAAACCAAGAGAGCGTGCACCAACGTTGTTAGGC 600
594 TATGACCCGAGATTTGATGCTGCTCCAAACCAAGAGAGCGTGCACCAACGTTGTTAGGC 653
601 ACAAAAGCAAGACAGTTGAGCAAAATCATCAAAAGCATCAAGCGCTTTAAGAACCCACC 660
654 ACAAAAGCAAGACAGTTGAGCAAAATCATCAAAAGCATCAAGCGCTTTAAGAACCCACC 713
661 AAAGTGACCAAGGTGTGCTGCTGAGACTGCCAACAGAGAGTATAGCAATTTGGTT 720
714 AAAGTGACCAAGGTGTGCTGCTGAGACTGCCAACAGAGAGTATAGCAATTTGGTT 773
721 GTAGGCTTTAATGACACATGAGAAATCTTGTGCTGCTGTTGAGACAGAAATAGGCTGAG 780
774 GTAGGCTTTAATGACACATGAGAAATCTTGTGCTGCTGTTGAGACAGAAATAGGCTGAG 833
781 ATTTTCTCTTCACTGTTATGCAATGCTGTGTGTGATGAAATGTTCTTTCAATTAAT 840
834 ATTTTCTCTTCACTGTTATGCAATGCTGTGTGTGATGAAATGTTCTTTCAATTAAT 893
841 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATTTGATCTTGGCAATGCCGAGAAACCT 900

|||||
894 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATTTGATCTTGGCAATCCCGAGAAACACT 953
901 TTGATTTGTGAGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGTTGAT 960
954 TTGATTTGTGAGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGTTGAT 1013
961 TTTTCTTGTGGGGCTGTGATTCAGGCCAATCTTATAGTTAGTTAACCAACATCTGGGAAAC 1020
1014 TTTCTTGTGGGGCTGTGATTCAGGCCAATCTTATAGTTAGTTAACCAACATCTGGGAAAC 1073
1021 AATGATGATGATATCTCTGGCTCCACCAACCTTCCGCTCCAGGAATCTCAAGAGC 1080
1074 AATGATGATGATATCTTGGCTCCACCAACCTTCCGCTCCAGGAATCTTCAAGAGC 1133
1081 AAGTTTGTGAGATATGCTCAACAGCAATGCAATCTTATAGCTGTGATGATGAT 1140
1134 AAGTTTGTGAGATATGCTCAACAGCAATGCAATCTTATAGCTGTGATGATGAT 1193
1141 GACCATGTTGTTGTTATTTAGTATGCTTACGTTAGGGGATAGCAAGAGCCATGAT 1200
1194 GACCATGTTGTTGTTATTTAGTATGCTTACGTTAGGGGATAGCAAGAGCCATGAT 1253
1201 GAGTACACTTCAAGATATTTCAATGAGGAGAAACACCATTTGTTTGGACACATGAT 1260
1254 GAGTACACTTCAAGATATTTCAATGAGGAGAAACACCATTTGTTTGGACACATGAT 1313
1261 GAGATTTCCCTTTTATGCTGCTCTTATTTATCTTGAATGTTGCTTCTTGTGAGCTGAC 1320
1314 GAGATTTCCCTTTTATGCTGCTCTTATTTATCTTGAATGTTGCTTCTTGTGAGCTGAC 1373
1321 ACTAGATTCAGTTTAAAGCTGAAATAGAGGAAATTTCACTCATTTCCACCAAGTTGCT 1380
1374 ACTAGATTCAGTTTAAAGCTGAAATAGAGGAAATTTCACTCATTTCCACCAAGTTGCT 1433
1381 ACCATTTCAAGCTATCTGACCAAGGCTCCTGTTTCCACCGGGATACCAAGTGTGAT 1440
1434 ACCATTTCAAGCTATCTGACCAAGGCTCCTGTTTCCACCGGGATACCAAGTGTGAT 1493
1441 GCATTTGCAAAAGCAGCTGCAATGCTGGAATAACATATGAGGCTTGTGTTGATTTGGCC 1500
1494 GCATTTGCAAAAGCAGCTGCAATGCTGGAATAACATATGAGGCTTGTGTTGATTTGGCC 1553
1501 CCAAGAAATTAATGATTTCTGAGTTACAGTGA 1533
1554 CCAAGAAATTAATGATTTCTGAGTTACAGTGA 1586
RESULT 3
BD075269 1533 bp DNA linear PAT 27-AUG-2002
LOCUS BD075269
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION BD075269.1 GI:22620872
VERSION UP 2001519665-A/4.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1533)
AUTHORS Hitz,W.D. and Sebastian,S.A.
TITLES Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid
JOURNAL Patent: JP 2001519665-A 4 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR33
PN JP 2001519665-A/4
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC
Strandedness: Double;

CC Topology: linear;
 CC Soybean plant producing seeds with reduced
 levels of raffinose
 CC and phytic acid saccharides
 FH Key Location/Qualifiers
 FT CDS 1..1533.

FEATURES
 source 1..1533
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN

Query Match 95.4%; Score 1462.6; DB 6; Length 1533;
 Best Local Similarity 97.1%; Pred. No. 0; Mismatches 44; Indels 0; Gaps 0;
 Matches 1489; Conservative 0;

1 ATGTTTCATCGAGATTTTAAAGTAGAGTCTCTAAATGTGAATGACACCGAGACTGAGATT 60
 1 ATGTTTCATCGAGATTTTAAAGTAGAGTCTCTAAATGTGAATGACACCGAGACTGAGATT 60
 61 CAGTCCTGTATCACTACGAAACCAACCGAATTGTTCAAGAGAACAGGAATGCACTAT 120
 61 CAGTCCTGTATCACTACGAAACCAACCGAATTGTTCAAGAGAACAGGAATGCACTAT 120
 121 CAGTGGATTTGCAAAACCCAAATCCGTCAACTCAATTTAAACCAACCCATGTTCCA 180
 121 CAGTGGATTTGCAAAACCCAAATCCGTCAACTCAATTTAAACCAACCCATGTTCCA 180
 121 CAGTGGATTTGCAAAACCCAAATCCGTCAACTCAATTTAAACCAACCCATGTTCCA 180
 181 AAATTTGGGGGTGATCTTGTGGGTTGGGGTGGAAACACGAGCTTACCCCTACCGGTGGT 240
 181 AAATTTGGGGGTGATCTTGTGGGTTGGGGTGGAAACACGAGCTTACCCCTACCGGTGGT 240
 181 AAATTTGGGGGTGATCTTGTGGGTTGGGGTGGAAACACGAGCTTACCCCTACCGGTGGT 240
 241 GTTATTGCTTAAAGAGAGACATTTTCAATGGGCTTACAAAGAGAACAAATTTCAACACCAT 300
 241 GTTATTGCTTAAAGAGAGACATTTTCAATGGGCTTACAAAGAGAACAAATTTCAACACCAT 300
 301 TACTTTGGGCTCCCTACCCCAAGCCTCAGCTATTCGAGTTGAGATTCCTCCAGGAGAGAGAA 360
 301 TACTTTGGGCTCCCTACCCCAAGCCTCAGCTATTCGAGTTGAGATTCCTCCAGGAGAGAGAA 360
 361 ATCTATGCCCATTTCAAGAGTGTGCTTCCAAATGTTAATCTCGACGACATTTGTTTGGG 420
 361 ATCTATGCCCATTTCAAGAGTGTGCTTCCAAATGTTAATCTCGACGACATTTGTTTGGG 420
 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAAAGGTGTTGAC 480
 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAAAGGTGTTGAC 480
 481 ATCGATTTGCAAGAGAGTGAAGGCTTACATGGAATCATGGTTCACCTCCCGGAATC 540
 481 ATCGATTTGCAAGAGAGTGAAGGCTTACATGGAATCATGGTTCACCTCCCGGAATC 540
 541 TAGCAGCCCGGATTTTCAATGCTGCTCAACCAAGAGAGCGCTGCCAACAACGTGATTAAGGCG 600
 541 TAGCAGCCCGGATTTTCAATGCTGCTCAACCAAGAGAGCGCTGCCAACAACGTGATTAAGGCG 600
 541 TAGCAGCCCGGATTTTCAATGCTGCTCAACCAAGAGAGCGCTGCCAACAACGTGATTAAGGCG 600
 541 TAGCAGCCCGGATTTTCAATGCTGCTCAACCAAGAGAGCGCTGCCAACAACGTGATTAAGGCG 600
 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAGACATCAAGCGCTTTAAGAGAGCCACC 660
 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAGACATCAAGCGCTTTAAGAGAGCCACC 660
 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAGACATCAAGCGCTTTAAGAGAGCCACC 660
 661 AAGATGGAACAAGGTGTGCTGCTGAGACTGCCAACAAGAGAGGTACAGTAATTTGGTT 720
 661 AAGATGGAACAAGGTGTGCTGCTGAGACTGCCAACAAGAGAGGTACAGTAATTTGGTT 720
 721 TAGAGCTTAATGACACATGAGAGATCTCTGGCTGCTGTGGACGAATATGAGGTGAG 780
 721 TAGAGCTTAATGACACATGAGAGATCTCTGGCTGCTGTGGACGAATATGAGGTGAG 780
 781 AATTCTCTTCCACCTTGTATGCAATGCTGCTGTGTGATGAAATGTTCTTTCAATTAAT 840
 781 AATTCTCTTCCACCTTGTATGCAATGCTGCTGTGTGATGAAATGTTCTTTCAATTAAT 840

841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATGATCTTGGCATGCGAGAGAACT 900
 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATGATCTTGGCATGCGAGAGAACT 900
 901 TTGATTTGATGAGATGACTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGGTAT 960
 901 TTGATTTGATGAGATGACTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGGTAT 960
 961 TTCTTGTGGGGGCTGGATATCAGCAACATCTATAGTTAGTTTCAACCATCTGGGAAAC 1020
 961 TTCTTGTGGGGGCTGGATATCAGCAACATCTATAGTTAGTTTCAACCATCTGGGAAAC 1020
 1021 AATGATGATGATGATCTCTCGGCTCCACAAACCTTCGCTCAAGAAATCTCCAAAGC 1080
 1021 AATGATGATGATGATCTCTCGGCTCCACAAACCTTCGCTCAAGAAATCTCCAAAGC 1080
 1081 AACGTTGTTGACATATGCTCAACAGCAATGCCATCTCTATGAGCCCTGGTGAACATCC 1140
 1081 AACGTTGTTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCCCTGGTGAACATCA 1140
 1141 GACCATGTTGTTGATTAATGATGATGCTTACGTAAGGAGATGACAGAGACCATGAT 1200
 1141 GACCATGTTGTTGATTAATGATGATGCTTACGTAAGGAGATGACAGAGACCATGAT 1200
 1201 GAGTACACTTCAGAGATATTCATGAGTGGTGAAGAACACCATTTGTTTGCACAAACATGT 1260
 1201 GAGTACACTTCAGAGATATTCATGAGTGGTGAAGAACACCATTTGTTTGCACAAACATGT 1260
 1261 GAGATTTCCCTTTTATGCTGCTCTATTAATCTTGAACCTTGTGCTTGTGAGCTGAGC 1320
 1261 GAGATTTCCCTTTTATGCTGCTCTATTAATCTTGAACCTTGTGCTTGTGAGCTGAGC 1320
 1321 ACTGAAATCCAGTTTAAAGCTGAATAAGAGGAAATTCACATCTATCCACCCAGTTGCT 1380
 1321 ACTGAAATCCAGTTTAAAGCTGAATAAGAGGAAATTCACATCTATCCACCCAGTTGCT 1380
 1381 ACCATTTCTACGATATCTGACCAAGGCTCTCTGTTCCACCGGCTACCAAGTGTGAT 1440
 1381 ACCATTTCTACGATATCTGACCAAGGCTCTCTGTTCCACCGGCTACCAAGTGTGAT 1440
 1441 GCATTGTCAAAAGCAGCGTGCATATGCTGGAACATTAATGAGGGCTTGTGATTTGGCC 1500
 1441 GCATTGTCAAAAGCAGCGTGCATATGCTGGAACATTAATGAGGGCTTGTGATTTGGCC 1500
 1501 CCAGAGATTAACATGATCTCGAGTACAAAGTGA 1533
 1501 CCAGAGATTAACATGATCTCGAGTACAAAGTGA 1533

RESULT 4
 AF293970 1791 bp mRNA linear PLN 03-MAY-2001
 LOCUS AF293970
 DEFINITION Glycine max myo-inositol-3-phosphate synthase (MIRS) mRNA, complete cds.
 ACCESSION AF293970
 VERSION AF293970.1 GI:13936690
 KEYWORDS
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 1791)
 REFERENCE
 1 Hegeman, C.E., Good, L.L. and Grabau, E.A.
 Expression of D-myo-inositol-3-phosphate synthase in soybean.
 Implications for phytic acid biosynthesis
 Plant Physiol. 125 (4), 1941-1948 (2001)
 JOURNAL MEDLINE
 PUBMED 21196082
 1129373
 2 (bases 1 to 1791)
 REFERENCE
 2 Hegeman, C.E., Good, L.L. and Grabau, E.A.

TITLE	Direct Submission
JOURNAL	Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed

FEATURES	Location/Qualifiers
source	1. .1791

ORIGIN

Query Match 95.0%; Score 1456.2; DB 8; Length 1791;

QY	1	ATGTCATCGAAGATTTTAAAGSTGAGAGTCTCTAATGTGAAGTACCGGAGACTGAGATT	60
Db	62	ATGTCATCGAAGATTTTAAAGSTGAGAGTCTCTAATGTGAAGTACCGGAGACTGAGATT	121
QY	61	CAGTCCGCTGTACAACCTACGAAACCGGAACTGTGTACGAGAAACGGAATGGACCAT	120
Db	122	CAGTCCGCTGTACAACCTACGAAACCGGAACTGTGTACGAGAAACGGAATGGACCAT	181
QY	121	CAGTGAATTGTCAAAACCCCAATCCGTCACATACCAATTTAAACCAACCCATGTTTCA	180
Db	182	CAGTGAATTGTCAAAACCCCAATCTGTCAAAATACGAATTTAAACCAACATCATGTTCT	241
QY	181	AAATTGGGGGTGATGCTTGTGGGTTGGGGGTGAAAACAACGGCTCTACCGGTGGT	240
Db	242	AAATTGGGGGTGATGCTTGTGGGTTGGGGGTGAAAACAACGGCTCTACCGGTGGT	301
QY	241	GTTATTGCTAAACAGAGAGACATTTCAATGGGCTACAAAGACAAAGATTCAACAGCCAT	300
Db	302	GTTATTGCTAAACAGAGAGGCGATTTCATGGGCTACAAAGACAAAGATTCAACAGCCAT	361
QY	301	TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATGCTTCACAGGAGAGGAA	360
Db	362	TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATGCTTCACAGGAGAGGAA	421
QY	361	ATCTATGCCCATTCGAAGACTGTGCTTCAATGGTTATCCTGAAGCATTTGTGTTGGG	420
Db	422	ATCTATGCCCATTCGAAGGCTGTCTTCAATGGTTAAACCTGAAGCATTTGTGTTGGG	481
QY	421	GGATGGGATATCAAGCAATGAACCTGTGGTGAATGCCATGGCAGAGGCAGAAAGTGTGGAC	480
Db	482	GGATGGGATATCAAGCAATGAACCTGTGGTGAATGCCATGGCAGAGGCAGAAAGTGTGGAC	541
QY	481	ATGATTTTGGAGAAAGAGTGAAGGCTTTCATGTGAATTCATGTTTCCAGTCCCGGAATC	540
Db	542	ATGATTTTGGAGAAAGAGTGAAGGCTTTCATGTGAATTCATGTTTCCAGTCCCGGAATC	601
QY	541	TACGACCCGGATTTCAATTGCTGCAACCAAGAGAGCGTGCACCAACATGTGATTAAAGGC	600

Db	602	TATGACCCGGATTTCATTGCTGCGCACCCAGAGAGCGTGCACCAAGTCATCAAGG3C	661
Qy	601	ACAAAGCAAGGCAAGTTCAGCAAAATCATCAAGACATCAAGAGCGTTTAAAGAACACC	660
Db	662	ACAAAGCAAGGCAAGTTCACAAATCATCAAAACACATCAAGGCGTTTAAAGAACCAAC	721
Qy	661	AAAGTGCAAGAGTGTGTTCTCTGTGACCTGCCAACACAGAGAGGTATAGCAATTTGGTT	720
Db	722	AAAGTGCAAGAGGTGTGTACTGTGTGACCTGCCAACACAGAGAGTACAGTAATTTGGTT	781
Qy	721	GTAGGCGCTTAATGACCAATGAGAAATCTCTGTGCTGTGAGACAGAAATAGAGGCTAG	780
Db	782	GTGGGCGCTTAATGACCAATGAGAAATCTCTGTGCTGTGTGACAGAAATAGAGGCTAG	841
Qy	781	ATTTCTCCCTTCACCTTGTATGCCATTGCTGTGTGATGAGAAATGTTCCTTTCAATAT	840
Db	842	ATTTCTCCCTTCACCTTGTATGCCATTGCTGTGTATGAGAAATGTTCCTTTCAATAT	901
Qy	841	GGAGCGCCTCAGAACCTTTTGTATCCAGGGGTGATATCTTTGCACTCGAGGAGAACCT	900
Db	902	GGAGCGCCTCAGAACCTTTTGTATCCAGGGGTGATATCTTTGCACTCGAGGAGAACCT	961
Qy	901	TTGATGTGTGAGATGACTTCAGAGTGTGTGACCAAAATGAATCTGTGTGGTGTAT	960
Db	962	TTGATGTGTGAGATGATTCCAAGTGTGTGACCAAAATGAATCTGTGTGGTGTAT	1021
Qy	961	TTTTTTTGGGGGGGTGATTCAGGCCAACATCTATATTTAGTTAACCAATCTGGGAAC	1022
Db	1022	TTTTTTTGGGGGGGTGATTCAGGCCAACATCTATATTTAGTTAACCAATCTGGGAAC	1081
Qy	1021	AATGATGATGATATCTCTCGGCTCCACAACTTCGCGTCCAGAGAAATCTCAGAGAC	1080
Db	1082	AATGATGATGATATCTTCGGGCTCCACAACTTCGCGTCCAGAGAAATCTCAGAGAC	1141
Qy	1081	AAGCTTGTGACCATATGCTGCACAGCAATGCCATCTCTATGAGCCTGTGTAACTCTCC	1140
Db	1142	AAGCTTGTGATGATATGCTGCACAGCAATGCCATCTCTATGAGCCTGTGTAACTCTCA	1201
Qy	1141	GACCATGTGTGTATTTAAGTATGTCCCTTACGTAGGGGATAGCAAGAGCGCATGGAT	1200
Db	1202	GACCATGTGTGTATTTAAGTATGTCCCTTACGTAGGGGACAGCAAGAGCGCATGGAT	1261
Qy	1201	GAGTACACTTCAGAGATATTCATGAGGTGAGAAAGAACCACTGTGTTTGCAACAACATGT	1260
Db	1262	GAGTACACTTCAGAGATATTCATGAGGTGAGAAAGAACCACTGTGTTTGCAACAACATGC	1321
Qy	1261	GAGGATTCCTGTTTAAAGTGTCCCTTAATCTTGGACCTTGGCTTCTTGTGAGACTGAGC	1320
Db	1322	GAGGATTCCTGTTTAAAGTGTCCCTTAATCTTGGACCTTGGCTTCTTGTGAGACTGAGC	1381
Qy	1321	ACTAGAAATCAGTTTAAAGGTGAAAAATGAGGGAAAAATTCACATCATTCACCCAGTTGCT	1380
Db	1382	ACTAGAAATCAGTTTAAAGGTGAAAAATGAGGGAAAAATTCACATCATTCACCCAGTTGCT	1441
Qy	1381	ACCATTTCTGAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGATACCAACAGTGTGAT	1440
Db	1442	ACCATTTCTGAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGATACCAACAGTGTGAT	1501
Qy	1441	GCAATTTCAAGAGCGCTGCAATGTGAGAAAACTAATGAGGCGTTGTGTGATTTGGCC	1500
Db	1502	GCAATTTCAAGAGCGGTGCAATGTGAGAAAACTAATGAGGCGTTGTGTGATTTGGCC	1561
Qy	1501	CCAGAGAAATACATGATTTCTCGAGTACAAAGTGA	1533
Db	1562	CCAGAGAAATACATGATTTCTCGAGTACAAAGTGA	1594

RESULT 5	LOCUS	DEFINITION
AB032073	AB032073	1931 bp mRNA linear Nicotiana glauca NpIMP1 mRNA for myo-inositol-1-phosphate synthase, complete cds.

ACCESSION AB032073
VERSION AB032073.1 GI:5834499
KEYWORDS myo-inositol-1-phosphate synthase.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 1931)
AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
TITLE Myo-inositol-1-phosphate synthase
JOURNAL Published Only in Database (1999)
AUTHORS 2 (bases 1 to 1931)
TITLE Hashimoto, A., Yamada, S. and Komori, T.
JOURNAL Direct Submission
Submitted (01-SEP-1999) Akiko Hashimoto, Japan Tobacco Inc., Plant
Breeding and Genetics Research Laboratory, 700 Higashibara, Iwata,
Toyoda-cho, Shizuoka 438-0802, Japan
(E-mail: Akiko.Hashimoto@pbgrl.jti.co.jp, Tel: +81-538-32-7116,
Fax: +81-538-33-6046)
FEATURES
source 1. 1931
Location/Qualifiers
/organism="Nicotiana paniculata"
/mol_type="mRNA"
/db_xref="taxon:62141"
1. 1931
/gene="NPINP1"
92. 1624
/gene="NPINP1"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="BAA84084.1"
/db_xref="GI:5834500"
/translation="MEIEFNPVSEPNVKTSEIHSVYDYOTTELVHDEKNGTYOMTV
KPKTVKFEKDVHVPKLGVMVGMGNSSTLGGVLANREGISMAIKDQKQANVE
GSUTASTIRVUSFNGEBIYAPFKSLIPRVNDDVPRGMDISMDLADMAKAVPD
IDLQKOLRYMESWVPLPETIDPDLIAAQSGRANNAVIGKEDIDILIKDIREFKE
NKKVDKVVLTWANTERSNVVGLNDTMEHFAVDREAEISSTLYALACILEEN
PFINSPQTFVPLGIDLAIKKNTLIGDDFKSCQTKMSVLDVLVAGIPTSIVS
YVHLNNGDMNLASOPTSRSEIKSNVDDVSSNALIYEGSEHDPVVVXYVYV
GDSKRAMDEYSEIFMGKNTVLAHNTGDSILAAPIILDVILAEISRIOLKKEG
GKFSPHFPVATLSTLTKAPLVFPPTPVVNLISKORAMLENTLRACVGLAPENNMLE
YK"

ORIGIN
Query Match 71.6%; Score 1097.8; DB 8; Length 1931;
Best Local Similarity 82.3%; Pred. No. 8.3e-289;
Matches 1261; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
1 ATGTTTCATCGAGATTATTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
2 AGTTTATGAGATTATTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 151
61 CAGTCCGTGTACACTACGAAACACCGAGACTTGTTCACGAGAACGAGAAATGCGACTT 120
152 CACTGTCTATGATTATCAAACTGAGTATGTTTCATGATGAGAAATAATGGGACATAT 211
121 CAGTGAATGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCAACCCATGTTTCA 180
212 CAATGACCGTCAAGCTTAAGACTGTCAATATGATTAAGATTAAGATTTATGTTTCCC 271
181 AAATGGGGGTGATGCTTGTGGGTGGGTGAGAAACAGGCTTACCCCTCAACCGGTGT 240
272 AAATTAGGGGTATATGCTTGTGGATGGGTGAGAAACATGTTCAACCTTGAACCGGTGT 331
241 GTTATTGCTTACAGAGAGACATTTATGAGGCTTACAAAGAACATTTCAACAGGCAT 300
332 GTTATTGCTTACAGAGAGAAATTTTCATGGGCCACAAAGATTAAGGTGCAACAGGCAT 391
301 TACTTTGGCTCCCTCAACCAAGCTCAGGCTATTCAGATTGAGATCTTCCAGGGAGAGAA 360
392 TACTTTGGCTCCCTTACTAGGCTTCTACTATTGAGTTGGGCTTTTCAATGAGAGAG 451

361 ATCTATGCCCATTCAGAGAGTCTGTTCCATATGTTAATCTGTACAGACATTTGTTGGG 420
452 ATCTATGCTCCATTTAAAGAGCTCTTCCATATGTTCAATCCAGATGAGTGTGTTGGA 511
421 GGATGGGATATCAGCAACATGAACTGGCTGATGCAATGGCCAGAGGCAAGGTGTTGAC 480
512 GGAATGGCATACAGGACATGAATTTTACAGATGCCATGGCCAGGGCTTAAGTATTTGAT 571
481 ATCGATTTGCAGAAAGCACTTGAAGGCTTTACATGAAATCATGTTTCACTCCCGGAATC 540
572 ATTGATCTACAAAGCAATGTAGAGGCTTACATGAAATCTATGTGTCACCTCCGTGATC 631
541 TAGACCCGGATTTTATGTTGCTGCCCAACCAAGAGAGCTGCCAACAGTATTAAGGC 600
632 TAGACCTCGATTTTATGTTGCTGCCCAACCAAGAGAGCTGCCAACAGTATTAAGGC 691
601 ACAAAGCAAGAGCAAGTTACAGCAATCATCAAGACATCAAGGGCTTAAGAGAGCCAC 660
692 ACCAAGAAAGAAACAAATGATCAATCATTTAAGATTTAAGGAGTTTAAGGAAAGAAC 751
661 AAAGTGACAAAGTGTGTTGCTGTGAGTCCCAACAGAGAGGTATAGCAATTTGGTT 720
752 AAAGTGACAAAGTGTGTTGCTGTGAGTCCCAACAGTAACTGAAGATACATATGTTGTT 811
721 GTAGGCTTTATGACATGACATGAGATCTTTGGCTGCTGTGAGACAGAAATGAGGCTGAG 780
812 GTTGGACTTAATGACATGATGAGAAACCTTTGCTGTGAGACAGAAATGAGGCTGAA 871
781 ATTTCTCTTCACCTTGTATGATGATGATGCTGTGTGATGAGAAATGTTCTCTTCTAAT 840
872 ATATCTCTTTCACCTTGTATGATGATGATGCTGTGTGATGAGAAATGTTCTCTTCTAAT 931
841 GGAAAGCTTCAGAAACATTTTGTACAGAGGCTGATGATCTTGTGATGAGAAAGTCT 900
932 GGAAGCCCCAGAAACATCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 991
901 TTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
992 TTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1051
961 TTTCTTGGGGGGGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1052 TTCTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1112 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1171
1081 AAGTTGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1172 AATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
1141 GACGATGTTGTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1232 GACGATGTTGTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
1201 GAGTACCTTACAGATATTTACAGGAGTGAAGAAACACCATTTGTTTCAACAAACATGT 1260
1292 GAGTACATCTGATATTTTACAGGAGTGAAGAAACACCATTTGTTTCAACAAACATGT 1351
1261 GAGATTCCTTTTACGTTGCTCTTATTTATCTTGAACCTTGTCTTGTGAGTGAAC 1320
1352 GAGGATTTACATTTTGTGCTCTTCAATTTATGATTTGTTGTTCTTGTGAGTGAAC 1411
1321 ACTAGATTCAGTTTAAAGTGAAGTGAAGGAAATTCATCTATTCACACCACTTGTCT 1380
1412 ACCCGATTTACGTTTAAAGTGAAGTGAAGGAAATTCATCTATTCACACCACTTGTCT 1471
1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTTTCAACCGGATGACCAAGTGTGAT 1440
1472 ACTATCTCTAGTATCTTACCAAGGCTCTCTGTTTCAACCGGATGACCAAGTGTGAT 1531
1441 GCATTGTCAAGCAGCGTGAATGCTGAGAAACATTAATGAGGCTTGTGTTGATTTGGCC 1500

Db 1532 GCACCTCAGAGAGAGGCGCAATGCTTGAGAACATTTGAGGGCTTGCTGAGACTTGA 1591
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTCAAGTGA 1533
Db 1592 CCAGAGAACACATGATTTCTGGAATACAAATGA 1624

RESULT 6
E27176
LOCUS E27176 1950 bp DNA linear PAT 18-JUN-2001
DEFINITION Novel INPS gene derived from nicotiana.
ACCESSION E27176
VERSION E27176.1 GI:13026394
KEYWORDS JP 1999187879-A/1.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 1950)
Shigehiro, Y. and Toshiyuki, K.
Novel INPS gene derived from nicotiana
Patent: JP 1999187879-A 1 13-JUL-1999;
JAPAN TOBACCO INC
OS Nicotiana paniculata
PN JP 1999187879-A/1
PD 13-JUL-1999
PR 26-DEC-1997 JP 1997359773
PI SHIGEHIRO YAMADA, TOSHIYUKI KOMORI
PC C12N15/09//C12N5/10, (C12N15/09, C12R1:91), C12N15/00, C12N5/00,
PC (C12N15/00, C12R1:91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 92..1703.
FEATURES
source Location/Qualifiers
1..1950
/organism="Nicotiana paniculata"
/mol_type="genomic DNA"
/db_xref="taxon:62141"

ORIGIN
Query Match 71.6%; Score 1097.8; DB 6; Length 1950;
Best Local Similarity 82.3%; Pred. No. 8.3e-289;
Matches 1861; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
Qy 1 ATGTCATGAGAAATTTAAGTAGAGAGTCCCTAATGTAAGTACCCGAGCTGAGATT 60
Db 92 ATGTTATGAGAAATTTAAGTAGAGAGTCCCTAATGTAAGTACCCGAGAGTGAATTT 151
Qy 61 CAGTCCGTGTACAACTACGAAACACCGAAGTTGTTACGAGAACAGAAATGGACCTAT 120
Db 152 CACTGTGTATGATTATCAAAACCACTGAGTTAGTTCAAGTAGAGAAATGGACATAT 211
Qy 121 CAGTGAATGTCAAAACCAATCCGTCAACTACCAATTTAAACCAACCAACCTGTTCA 180
Db 212 CAATGACCGTCAAGGCTTAAGCTGCAATATGATTCAGAGCTGATGATGTTCC 271
Qy 181 AAATTGGGGGTGATGTTGGGTGGGGTGGAAACAAGGCTTACCTCAACCGTGT 240
Db 272 AAATTAGGGGTATGCTTGTGTGATGGGTGGAACAAATGTTCAACCTTGAACCGTGT 331
Qy 241 GTTATGCTAACAGAGACATTTTCATGGGCTACAAAGACAAAGATTCAACAGCAAT 300
Db 332 GTTATGCTAACAGAGAGATTTTCATGGGCTACAAAGATTAAGTGCACAAAGCAAT 391
Qy 301 TACTTTGGCTCCCTCAACCAAGGCTCAGTATTGAGATTCTTCCAGGGAGAGAA 360
Db 392 TACTTTGGCTCTTACTCAGAGCTTCTACTATTGAGGTGGGTCTTTCAATGAGAGAG 451
Qy 361 ACTATAGCCCATTAAGAGTCTGCTTCAATGTTAATCTGACACATTTGTTGGG 420

Db 452 ATCTATGCTTCATTTAAAGCCTCCTTCCATGCTCAATCCAGATGACGTATGTTGA 511
Qy 421 GGATGGATATACGACAACTGAACCTGGCTGATGCCAGGCGCAAGGTGTTGAC 480
Db 512 GGATGGACATCAGGACATGAAATTTAGCATGTCATGGCAGGGCTAAGGATTTGAT 571
Qy 481 ATCGATTTGACAGAGAGCTTGAAGCCTTACATGAAATCCATGCTCCCGGAATC 540
Db 572 ATTGATCTACAAAGAGTTGAGGCTTACATGAAATCTATGCTCCACTCCCTGATTC 631
Qy 541 TAGACCCCGATTTTCATGCTCTCCCAACGAGAGAGCGTGCACAACTGATTAAGGCG 600
Db 632 TATGACCTCGATTTTCATGCTCTCTACCAAGGCTACGTCGCAACACTGATCAAGGA 691
Qy 601 ACAAGCAAGACAGTGTGATGCAAAATCATCAAGATCAAGAGCGTTTAAGAGGACCC 660
Db 692 ACCAAGAAAGAACAAATGATCATTAAGATATTAAGGAATTAAGAAAGAAC 751
Qy 661 AAAGTGAACAGTGTGCTCTGTTGAGTGCACACAGAGATTAAGCAATTTGTT 720
Db 752 AAAGTGAACAGTGTGATGATTTGAGACTGCTTAACAGTAAGTAATGTTGTT 811
Qy 721 GTAGGCTTAATGACACCATGAGAAATCTTGGCTGTGAGACAAATGAGGCTGAG 780
Db 812 GTTGAATTAATGACATATGAAACCTCTTGGCTGTGAGACAAATGAGGCTGAA 871
Qy 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGAGAAATGTTCTTTCATTAAT 840
Db 872 ATATGCTCTTCCACCTTGTATGCTATGCTGATTCGATTCGAAATGTCCTTTATTAAT 931
Qy 841 GGAAGGCTCAGAACACTTTTGTACAGGCTGATTTGATCTTGGCTCCAGAGAACT 900
Db 932 GGAAGGCTCAGAACACTTTTGTCCAGGCTCATTTGATTTGGCAATCAAGAAAGACAA 991
Qy 901 TTGATTTGGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTTGTTGAT 960
Db 992 TTGATTTGGAGATGACTTCAAGAGTGTGACCAAAATGAAATGCAAGTCTGTTGAT 1051
Qy 961 TTTCTTGGGGGCTGTATCAAGCAACATCTATAGTTATCAACCATCTGGGAAAC 1020
Db 1052 TTCTTGTGAGAGTGTATTAAGCAACATCAATGTGAGTACCAACATTTGGGTAAC 1111
Qy 1021 AATGATGATTAATCTCTGCTCCCAACACTTCCGCTCCAGGAATCTTCAAGAGC 1080
Db 1112 AATGATGATTAATCTCTGCTCCCAACACTTCCGCTCCAGGAATCTTCAAGAGC 1171
Qy 1081 AAGTGTGAGATATGCTCAAGCAATGCAATGCTCTATGAGCTGTGTAACATGCC 1140
Db 1172 AATGTTGTGATGACATGTTTCAAGCAATGCAATGCTCTTATGAGCTGTGAGACCTT 1231
Qy 1141 GACCATGTTGTTATTATTAAGTATGCTTACGTAGGGGATAGCAAGAGCCATGAT 1200
Db 1232 GACCATGTTGTTATTATTAAGTATGCTTATGTTGAGAGACAGCAAGGGCAATGAT 1291
Qy 1201 GAGTACACTTCAAGATATTCATGAGTGAAGAAACACATTTGTTGCAACACATGT 1260
Db 1292 GAGTACACTTCAAGATATTCATGAGTGAAGAAACACATTTGTTGCAACATGT 1351
Qy 1261 GAGGATTCCTTTTGAAGCTGCTCTATTAATCTTGGACTTGTGCTTGTGAGGCTGAGC 1320
Db 1352 GAGGATTCCTTTTGAAGCTGCTCTATTAATCTTGGACTTGTGCTTGTGAGGCTGAGC 1411
Qy 1321 ACTAGATTCAGTTTAAGCTGAAGAAATGAGGAAATTCACCTCAATCCACAGTTGCT 1380
Db 1412 ACCGCAATTCAGCTCAAGCTGAAGAGAGGTATGTTCACTCTTCAACCCCGTGGCT 1471
Qy 1381 ACCATTCAGCTATCTGACCAAGGCTCTGTTTCCACGGGATCACCAAGTGTGAT 1440
Db 1472 ACTATCTCAGCTATCTTACCAAGGCTCTCTGTTTCCACCAAGGATCACTAGTGTGAT 1531
Qy 1441 GCAATGTCAAGCAGGCTGCAATGCTGAAACATTAAGAGGCTTGTGTTGATGGCC 1500

Db	1532	GGACCTCTGAACGACGAGGCATTCGTTGAGAACACATTTGAGGCGCTTGCTTGACCTTGCA	1591
Qy	1501	CCAGAGAAATACATGATTTCTCGAGTACAACTGCA	1533
Db	1592	CCAGAGAAACAACATGATTTCTCGAAATACAAATGA	1624
RESULT 7			
AB009881			
LOCUS			
DEFINITION	AB009881	1959 bp	mRNA
ACCESSION	AB009881		linear
VERSION	AB009881.1	GI:8096265	
KEYWORDS	myo-inositol 1-phosphate synthase.		
SOURCE	Nicotiana tabacum (common tobacco)		
ORGANISM	Nicotiana tabacum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Lamiales; Solanales; Solanaceae; Nicotiana.		
AUTHORS	Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.		
TITLE	1 (sites)		
JOURNAL	Screening of wound-responsive genes identifies an immediate-early		
MEDLINE	expressed gene encoding a highly charged protein in mechanically		
PUBMED	wounded tobacco plants		
REFERENCE	Plant Cell Physiol. 41 (6), 664-691 (2000)		
AUTHORS	2 (bases 1 to 1959)		
TITLE	Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.		
JOURNAL	Direct Submission		
MEDLINE	Submitted (18-DEC-1997) Kojiro Hara, Nara Institute of Science and		
PUBMED	Technology, Plant Molecular Breeding; Takayama 8916-5, Ikoma, Nara		
REFERENCE	630-0101, Japan (E-mail:k-hara@bs.aist-nara.ac.jp,		
AUTHORS	Tel:+81-743-72-5653(ex.5653), Fax:+81-743-72-5659)		
TITLE	location/Qualifiers		
FEATURES	1..1959		
source	/organism="Nicotiana tabacum"		
	/mol_type="mRNA"		
	/db_xref="taxon:4097"		
	/tissue_type="leaf"		
	94..1676		
	/codon_start=1		
	/product="myo-inositol 1-phosphate synthase"		
	/protein_id="BA095788.1"		
	/db_xref="GI:8096266"		
	/translation="MEINPKYESPWVYKTESEHSVYDYOTTELVEHEKNGTYOMTYR		
	KPKTYAEKPKDVAHPKLGWLVGMGNGNSTLNGVIANEGISMAITKDVQOANFR		
	GSULTASTIRGSPFGEETIYAPFKSLPMWNPDDVYREGMISGMNLADMARKVDP		
	IDLQKVLPRVMSWVPLPGIYDPDFLIANQSGRANVYKTKRKQIDQIADIEFEN		
	KNKDVLLWNTANTERSNVNVGANDMENLFSVSDNEAEISPTLYAIACTLEEV		
	YFINSPOINTFVPGILDLAIKRNLTIGDDPKSGQTKKSVLDFLVAGIKPISVSG		
	VNHLNNDGNNLWAPQTERSKESIKSNVDDMSGNATLYERGHPDHVVITKVPYVA		
	GDSKRAMPEYSEIFMGSKNTIVHNTCEDSLIAPITLDIVILAELSTRQLKAGEE		
	GKFHSFHPVATITLSTYTKAPLVPPEPTPVNALSKORAMLENITLACVGLAPENNMTLE		
	YK"		
	1959		
	/note="15 a nucleotides"		
polya_site			
ORIGIN			
Query Match	71.3%;	Score 1093;	DB 8;
Best Local Similarity	82.1%;	Pred. No. 1.7e-287;	Length 1959;
Matches 1258;	Conservative	0;	Mismatches 275;
		Indels	0;
		Gaps	0;
Qy	1	AGTTCATCGAATAATTTTAAGGTAAAGAGTCTCTAATGGAAGTACACGAGACAGAGT	60
Db	94	AGGTTATATGAACAACCTTAAAGGTGAGAGCCCAACGTTAAGTACACGAAAGTGAATT	153
Qy	61	CAGTCGCGTGTAACTACGAGAAACCAACGAACTTTGTCACGAGAACAGAGACGACCTAT	120
Db	154	CAGTCGCTCATGATTTATTAACCACTAGATTGATTCATGAGAGAGAAAATGGGACTTAC	213
Qy	121	CAGTGAATGTCAAACCAAACTCGGTCAACTACCAATTTAAACCAACCAACCATGTTCCA	180

Db	214	CAATGACCTGCAAGCCTTAAGACTGCTCAAAATATGAGTTCAAGACTGATGTTCAATGTTCC	273
QY	181	AAATTTGGGGGTGATGCTTGTGGGTTGGGGGTGAAAACAACGGCTTCACTCCACCGGTGCT	240
Db	274	AAATTAGGGGTATTATGCTTGTGGATGGGGGTGAAAACAATAGTTTCAACCTTGACCGGTGCT	333
QY	241	GTTATTTGCTACAGAGAGGACATTTCAATGGGCTTCAAAAGGACAGATTTCAACAAGCCAAAT	300
Db	334	GTTATTTGCTAACAGAGAGGAAATTTCAATGGGCTCAAAAGTAAGGTGCAACAAGCCAAAT	393
QY	301	TACTTTGGCTCCCTCAACCAGCCTCAGCTTTTGAGTTGGATCTTTCAGGGAGAGGAA	360
Db	394	TACTTTGGCTCTTTCACAGGCTTTCACATTTGAGTTGGGTCTTTCAATGAGAAAG	453
QY	361	ATCTATGCCCCATTCAAGAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGG	420
Db	454	ATCTATGCTCATTTTAAAGCCTCTTCCAAATGCTCAATCAGATGACGTATGTTTGG	513
QY	421	GGAATGGATATCAGACACATGACCTGCTGATGCTCATGGCAGGGCAGAAAGTGTTCAC	480
Db	514	GGAATGGGACATCACACGGGATTAATTTGGCAGATCCCATGGCAAGGGCTTAAGATTTGAT	573
QY	481	ATCGATTTGCAAGAGCAGTTGAGGCTTACATGGAATCCATGTTCCACTCCCGGGAATC	540
Db	574	ATTGATCTCAAAAAGCAGTTGAGGCTTCACTGAAATCTATGTTCCACTACCTGGTATC	633
QY	541	TACGACCCGGAATTTCAATTGCTGCCAACCAAGAGAGCGGTGCCAAACAACGTATTTAAGGC	600
Db	634	TATGATCTGATTTCAATTGCTGCTGATACCAAGAGGTACAGTGCCAACAAGTATCAAGA	693
QY	601	ACAAAGCAGAGCAAGTTCAAGCAATCATCAAGAATCAAGGCGTTTAAAGAAAGCCACC	660
Db	694	ACCAAGAAAGAAACAAATTTGATCAATCATCTTAAGATATTTAGGAGTTTAAAGAAAGAAC	753
QY	661	AAAGTGACAGAGGTGGTGTCTGTGGACATGCAAGAGAGGTATAGCAATTTGGATT	720
Db	754	AAAGTTGACAGAGGTGATGTTGTGGATCTCTAACCTGAAATATACAGTAAATGTGTT	813
QY	721	GTAGGCTTAAATGACACCATGAGAAATCTCTTGGCTGCTGTGACAGAAATGAGGCTGAG	780
Db	814	GTTGGACTTAATGATACTATGAGAAATCTCTTGGCTTCTGTGGACAGAAATGAGGCTGAA	873
QY	781	ATTTCTCTTCCACTGCTGTATGCCATTTGCTCTGTGTATGAAATATGTTCTTTCAATTAAT	840
Db	874	ATATCTCTTCCACTTGTGTATGCTATGCTCCGCAATCTTGTGAAATATGCTCTTTATTAAT	933
QY	841	GGAAGCCCTCAGAACATTTTGTACAGGGCTGATGATCTTGGCATGCGAGGAACACT	900
Db	934	GGAAGCCCAAGAACACCTTTGTGCAAGGCTCAATGATTTGGCATCAAGAGAACACA	993
QY	901	TTGATTTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGGTTGAT	960
Db	994	CTGATTTGGTGTGATGACTTTAAAGATGTGTAAACAAAATGAAGTCACTGTCTGTTGAT	1053
QY	961	TTTCTTTGGGGGCTGGTATCAAGCCCAATCTATAGTTAGTTAACACATCTTGGGAAAC	1020
Db	1054	TTCTCTTTGGAGCTGGTATTAAGCCCAATCAATTTGTGAGCTTCAACATTTTGGGTAAAC	1113
QY	1021	AATATGTTAATGAAATCTCTCGGCTCCAAACCTTCGCTCCAGAGAAATCTCCAAAGAC	1080
Db	1114	AATATGTTAATGAATCTGTCTGCTCCCAACCTTCGCTCCAAAGGAGATCTCAAAAAGT	1173
QY	1081	AACGTTGTTGACATATGCTCAACAGCAATGCAATCCCTATGAGCCTGGTGAACATCCC	1140
Db	1174	AATGTTGTTGATGACATGTTTCAAGCAATGCCATTTCTTATGAGCCTGGAGAGACCTCT	1233
QY	1141	GACCAATGTTGTTGTTAATTAAGTATGTGCTTACGTAGGGGATAGCAAGAGACCATGAT	1200
Db	1234	GACCAATGTTGTTGTTAATTAAGTATGTGCTATATGTGGGAGACAGCAAAAGGGCAATGAT	1293
QY	1201	GAGTACATTTAGAGATATTCATGGGTGGAAGAACACATTTGTTTTCACAAACATGTT	1260

Db 1294 GAGTACACATCTGAGATTTTCATGAGGAGAAAGAACACATGTTTTCACAACTATGTT 1353
Qy 1261 GAGGATTCCTTTTACTGCTGCTTATTTATCTTGAGACTGGTCTTTCTGTGTGAGCTGAGC 1320
Db 1354 GAGGATTCCTTTTGGCTGCTGCTCAATATCTTGATTTGGTCTTTCTGTGTGAGCTGAGT 1413
Qy 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAGAAATTCACACTATTCACCCAGTTGGCT 1380
Db 1414 ACCCGATTCAGCTCAAGCTGAGAGAGAGGTAAGTTCCACTCTCCACCCCGTGGCT 1473
Qy 1381 ACCATTCAGCTATCTGACCAAGGCTCTCTGTGTTCACCGGGTAACCCAGTGTGAAT 1440
Db 1474 ACTATCTCAGCTACCTTACCAAGGCTCTCTGTGTACACAGGTACACCAAGTGTGAAT 1533
Qy 1441 GCATTGTCAAGGAGGTCGAATGCTGGAATACATATAGAGGCTTGTGTGATTTGGCC 1500
Db 1534 GCATCTCAAGAGAGGGGCAATGCTTGAGAACATATGAGGGCTGTGTGACTTGCA 1593
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1533
Db 1594 CCAGAGAACACATGATTTCTGGAATACAAATGA 1626
RESULT 8
BT013759 1954 bp mRNA linear PLIN 11-MAY-2004
LOCUS Lycopersicon esculentum clone 132642P, mRNA sequence.
DEFINITION BT013759
ACCESSION BT013759.1 GI:47105174
VERSION FLI_CDNA.
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1954)
Kirkness, B.F., Wang, W. and Vazeille, A.
REFERENCE Direct Submission
AUTHORS Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
TITLE Medical Center Drive, Rockville, MD 20850, USA
JOURNAL Location/Qualifiers
FEATURES
source 1..1954
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132642P"
/cbase_type="fruit"
/note="TZXB41"

ORIGIN

Query Match 69.7%; Score 1069; DB 8; Length 1954;
Best Local Similarity 81.1%; Pred. No. 6.3e-281;
Matches 1243; Conservative 0; Mismatches 290; Indels 0; Gaps 0;
Qy 1 ATGTCATCGAGAAATTTAAGTAGAGTCTTAATGTGAATATACACCGAGCTGAGATT 60
Db 107 ATGTTATTTAGAACATTTTAAGTTTGAAGCCCAATGTTAAGTACACAGAGGTAGATT 166
Qy 61 CAGTCGCTGTAACAACCAACCAACCACTGTGTCAAGAGACAGATGSCACTAT 120
Db 167 CACTCTGTGTATGATTAACAACCACTGAGTTGTTCAATGAGGAAAAAATGGACATTT 226
Qy 121 CAGTGGATTTGCAAAACCAATCCGTCAACTACCAATTTAAACCAACACCATTTTCA 180
Db 227 CAATGCACTGTAAACCAAAACCTGCAAAATGATTCAAAACTGATGTTGATTTCCA 286
Qy 181 AAATGGGGGTGATGCTGTGGGTTGGGGTGAACAACGCTTACCTCAACCGGTGT 240
Db 287 AAATTAAGAGTTATCTGTGGATGGAGTGAACAACATGTTGACCTTGACCGAGGT 346
Qy 241 GTTATTTGCTAACAGAGAGACATTTCAATGGGCTACAAAGACAAATTTCAACCAAT 300
Db 347 GTTATTTGCTAATAGAGAGAAATTTCAATGGGCTACAAAGATTAAGTGCAGAACCAAT 406

Qy 301 TACTTGGCTCCCTCACCCAGACCTCAGCTATTTGAGTTGATTCCTTCCAGGAGAGAA 360
Db 407 TACTTGGCTCTTACACAGAGCTCTTACTATTTGAGTTGGCTTTTCAATGAGAGAG 466
Qy 361 ATCTATGCCCAATTCAGAGTCTGCTTCAATGTTTATCTTGACGACATATGTTGGG 420
Db 467 ATCTATGCTCCCTTTAAGAGCATCTTCCATGTTATTCAGATATGATTTTGA 526
Qy 421 GATGGGATATACGAACATGAACCTGGCTGATGCCATGGCCAGAGGCAAGTGTGAC 480
Db 527 GATGGGACATCAGCAACATGAATTTGGAGATGCCATGGCAAGAGCTAAAGTTTGGAC 586
Qy 481 ATCGATTTGCAAGAGAGTGAAGGCTTACATGAAATCCATGTTTCACTCCCGGAATC 540
Db 587 ATTGATCTTCAATGAATGACATGAGGCTGTACATGGAATCCATGGTCCCACTGCTGTATC 646
Qy 541 TAGAACCCGGAATTTCAATGCTGCAACCAAGAGAGCGGCCAACAACTGATTTAGGGC 600
Db 647 TAGAACCTGATTTTCAATGCTGCAACCAAGGCTCAGTGTACATATCTCAAGGA 706
Qy 601 ACAAGCAAGACCAAGTTCAGCAAAATCATCAAGACATCAGCGTTTAAGAGACCAAC 660
Db 707 ACCAAGAAAGAACAAATGATGATCAAAATATTAAGATATTAGGAGTTTAAGGAGAACAC 766
Qy 661 AAAGTGAACAGTGTGTTCTGTGAGTGTCCCAACACAGAGATATGCAATTTGGTT 720
Db 767 AAAGTGAACAGGTTGTGTCTGTGAGTGTCCCAACACATGAAGATACAGCATGTGGCT 826
Qy 721 GTAGGCTTAATGACCACTGAGAAATCTTGGCTGCTGTGACGAAATGAGGTGAG 780
Db 827 GTTGGCTTAATGATACATGAGAAACCTTTGCTTCTGTGATGAATGAGGTGAA 886
Qy 781 ATTTCTCTTCCACCTTGATGATCCATGCTGTGTGATGAAGAAATGTTCTTTCAATAT 840
Db 887 ATATCTCTTCCACCTTGATGATCTTGTGTATTTCTTGAAGATGATACCTTCAATAT 946
Qy 841 GGAAGCCTCAGAACACTTTGTATCAAGGGCTGATTTGATCTTGCATGCGAGAACACT 900
Db 947 GGAAGCCTCAGAACAAATCTTTGTGCTCAGGCTCATTTGATGATGATGAGAGAACACT 1006
Qy 901 TTGATTTGGTGAAGTGAAGTCAAGAGTGTGACCAAAATGAATGTGTGTTGAT 960
Db 1007 TTGATTTGGTGAAGTGAAGTCAAGAGTGTGACCAAAATGAATGATGATGATGAT 1066
Qy 961 TTTCTTTGGGGGCTGTATCAAGCCCAATCTATAGTTAGTTAACAACCATCTGGGAAC 1020
Db 1067 TTCTTTGGGAGCTGTATTAAGCCACATCAATCGAGCTACACCATTTGGGTAAAC 1126
Qy 1021 AATGATGTATGATCTCTGGCTTCAACAACTTCCGCTCCAGAGAAATCTCAAGAGC 1080
Db 1127 AATGATGTATGATCTCTGGCTTCAACAACTTCCGCTCCAGAGAAATCTCAAGAGC 1186
Qy 1081 AAGCTTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1187 AAGCTTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246
Qy 1141 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1247 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
Qy 1201 GAGTACATTCAGAGATATTCAGAGGTTGAAGAAACACATTTGTTTTCACAAACATGT 1260
Db 1307 GAGTACATTCAGAGATATTCAGAGGTTGAAGAAACACATTTGTTTTCACAAACATGT 1366
Qy 1261 GAGATTCCTTTAGCTGCTCTTATTAATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1367 GAGATTCCTTTAGCTGCTCTTATTAATGATGATGATGATGATGATGATGATGATGAT 1426
Qy 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCACCTCAATTCACCCAGTTGCT 1380
Db 1427 ACCCGATTCAGCTCAAGCTGAAAGAGGATTAAGTTCACTCTTCCATCCCTGGGCT 1486

QY	1381	ACCAATTCTCAGCTACTCTGACCAAGAAGCTCTCTGGTTCCACCGGGGTACACAGTGGTAAT	1440
Db	1487	ACCAATCCCTCACTATTATCTTACCAAGGCTCTTTGGTACCCACAGGTACACAGTGGTAAT	1546
QY	1441	GCATTGTCTAAGCAGCGCTGCAATCTGGGAAATAATATGAGGGCTTGTGTGATTGGCC	1500
Db	1547	GCCCTTTCAAAGCAGAGGGCAATGCTCTGTAACATATTAGAGGCTTGTGTGACTGGCA	1606
QY	1501	CCAGAGAATAACATGATTCTCGAGTACAGTGA	1533
Db	1607	CCAGAGAAACAACATGATCTTGGAAATACAAATGA	1639
RESULT 9			
LOCUS	BT013505	1959 bp	mRNA linear PLN 11-MAY-2004
DEFINITION	Lycopersicon esculentum clone 132193F, mRNA sequence.		
LOCUS	BT013505		
ACCESSION	BT013505.1	GI:47104920	
KEYWORDS	FLI cDNA.		
SOURCE	Lycopersicon esculentum (tomato)		
ORGANISM	Lycopersicon esculentum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 1959)		
AUTHORS	Kirkness, E.F., Wang, W. and Vazeille, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-MAY-2004) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA		
FEATURES			
Source	Location/Qualifiers		
	1..1959		
	/organism="Lycopersicon esculentum"		
	/mol_type="mRNA"		
	/db_xref="taxon:4081"		
	/clone="132193F"		
	/tissue_type="fruit"		
	/note="TZB51"		
ORIGIN			
Query Match	69.2%;	Score 1061;	DB 8; Length 1959;
Best Local Similarity	80.8%;	Pred. No. 9.7e-279;	
Matches 1238; Conservative	0;	Mismatches 295;	Indels 0; Gaps 0;
QY	1	ATGTTTCATCGAGAATTTTAAGGTAGAGAGTCTCTAATGTGAAGTACCGGAGCTGAGATT	60
Db	77	ATGTTTATGAAAATTTTAAGGTGAAAAGCCCAATGTGAAGTATATTGAGATGAATTT	136
QY	61	CAGTCCGTGTACACTAGAAACCAACCGAATCTTGTTCACGAGAAAGAGAAATGGCACTAT	120
Db	137	CATTCTGTGTATGATTATGAAACCAAGAGCTGTGTCATGAAAGAGAAATGGAATCTTAT	196
QY	121	CAGTGAATGTGCAAAACCCAAATCCGTCAACTACCAATTTAAACCAACACCATGTTCCA	180
Db	197	CAATGATTTGTTAAAGCTTAAACTGTCAATATGAAATTTAAACTGATACCATGTGCA	256
QY	181	AAATTTGGGGGTGATCTGTGGGTTGGGGTGGGAGAAACAAGCTTACCTCAACGGTGT	240
Db	257	AAATTTGGGGGTATGCTTGTGGATGGGAGGAAACAATGTTCAACATTGACTGGAAGT	316
QY	241	GTTATTGTGCTAACAGAGAGACATTTCAATGGGCTACAAAGAGAAAGAAATTCACAAACCAT	300
Db	317	GTTATTGCGCAATCGAAGAAAGAAATTTCAATGGGCAAGAAAGAAAGATGCAACAACCTCAT	376
QY	301	TACTTTGGTCCCTCAACCAAGCTCAGCTATTTCGAGTTGAATCTTCCAGGAGAGGAA	360
Db	377	TATTTTGGGTCTCTTACTAGGCAATCAACATTCAGATTGGGCTTTTCAATGGCGAAGG	436
QY	361	ATCTATGCCCATTCGAAGAGTCTGCTTCCATAGTTTAATCTCGACGACATTGTGTTGGG	420
Db	437	ATCTATGACACCTTTCAAAAGCCTCCTTCCACATGATTAACCAAGAGATGATGATTTTGG	496
QY	421	GGATGGGATATACGACCAATGAACCTGGCTGATGCCATGCGCAGGCAAGGCAAAAGGTGTTGAC	480

Db		497	GGATGGGACATTTAGCAACATGAAATTTGGCAGATGCTATGGTCACGGGCTAAGATTTCGAA	556
Qy		481	ATCGATTTCGAAAGCAGTTGAGGCCCTTACATGGAATCCATGGTTCCATCCCGGGAATC	540
Db		557	GTTGATCTGGAAAAAGCAGCTGAGGCCCTCATATGAAATCCATGGTTCCCTTCGGTATC	616
Qy		541	TACGACCCGGATTTCAATTGCGGCCAACCAAGAGAGGGTGCACAAACGTGATTAAGGGC	600
Db		617	TATAGACCTGACTTCATTGGGGGCTAACCAAGAGACAGTGCACAAACGTATCAAGGA	676
Qy		601	ACAAAGCAAGCAGAGTTCAGCAAAATCATCAAGACATCAAGGCGTTTAAGAAACCAAC	660
Db		677	ACCAAGAAAGACAGTTGAAACAAATGTTTAAGATATATAGGAGGTTCAAGAGAAAGAC	736
Qy		661	AAAGTGAACAAAGTGGTTCCTGTGGACTGCGCAACACAGAGAGTATACCAATTTGGT	720
Db		737	AAGGTAGCAAGATAGGTGTTCTATGAGCTGCACAAACGAAAGGTACAGTAAATGEGTT	796
Qy		721	GTAGGCGCTTAATGACACCATGAGAAATCTGTGGCTGCTGTGGAACAGAAATGAGGCTGAG	780
Db		797	GTTGGCGCTTAATGACACCATGAGAAACCTTTAGCTGCTGTGGAATAGAAATGAGGCTGAA	856
Qy		781	ATTTCTCCCTTCACCTTGATGACCATTCCTGTGTGATGAGAAATGTTCTTCATTAAAT	840
Db		857	ATATCTCCTTCATATGACATGCATATGCTTGATATATGAAATATGCGCTTCATCAAC	916
Qy		841	GGAAAGCCTCAGAACACTTTTGTATCCAGGGGTGATGATCTTGGCATCCGAGGAACAAT	900
Db		917	GGAAAGCCTCAAAACACTTTTGTTCACAGGCTCTTATTAATTTGGCCATTAAGAGAAACAT	976
Qy		901	TTGATTTGGTGAAGATGACTTCAAGAGAGGTGAGACCAAAATGAAATCTGTGTGGTAT	960
Db		977	TTAATTTGGTGTATGACTTTTAAGAGTGTCAAAACAGATGAAGTCCGTGTGTGAT	1036
Qy		961	TTTCTTGGGGGGGTGGTATCAAGCCAAACATCTATAGTTAACAACATCTGGGAAAC	1020
Db		1037	TTCTGTGTGGAGCCGGATTAAGCCAAACGTCAATATGAGCTACAAATCTTGGGTAAAC	1096
Qy		1021	AATGATGGTAAATCTCTCGGCTCCACAAACCTTCCGCTCCAAAGAAATCTCCAAAGAC	1080
Db		1097	AATGATGAAATGAATCTTTCTGCTCCCAAAACCTTCGGGTCTAAGGAGATCTCAAAAGT	1156
Qy		1081	AACGTTGTTGACAGATAGGTGCAACAGGAAGCCATCCCTATATGAGCGCTGGTGAACATCCC	1140
Db		1157	AATGTTGTATGACATAGGTGTCTAGCAACCCATCTCTATGAAATCTGCGGACACCTT	1216
Qy		1141	GACCATGTTGTGTTAATAGTATATGCTTACGCTAAGGGGATAGCAAGAGCATGAT	1200
Db		1217	GACCATGTTGTGTATCAATATATGTTCCATATGTGAGACACAGAGGGCAATGGAT	1276
Qy		1201	GAGTACACTTCAGAGATATTCATGGGTGGAAGAACACATTTGTTTGGACAAACATATGT	1260
Db		1277	GAGTACATGTCAAGATATTTCAATGGGCGGAAGAGCACTAATAGTTATGCAAAACATGTT	1336
Qy		1261	GAGATTCCTCTTTAGCTGCTCCTAATATCTTGAACCTTGCTCTTCTGTGAGGTGAGC	1320
Db		1337	GAGGACTCTCTTTGGCAGCTCCATTTATCTTGAATTTGTCCTTCTTCGCTGAACCTCAGC	1396
Qy		1321	ACTAGAAATCCAGTTTAAAGCTGAATAATGAGGAAATTCACATCTATCCACCAGTTGCT	1380
Db		1397	ACTGCAATTCAGCTCAAAAGCTGAAGGGGAGGGGAAGTTCCATCTCTTCAACCTGTGGCT	1456
Qy		1381	ACCATTCTCAGCTATCTGACCAAGGCTCTTGTGTTCCACCGGGTACACCATGTGTAAAT	1440
Db		1457	ACTATTTCAAGTACCTCAACCAAGGCTCCCTGGTATACACCAAGGTACTCCAGTGTGTAAT	1516
Qy		1441	GCATTGTCAAAGCAGCGTGAATGTGGAATAAATAAGAGGGCTGTGTTGGAATGGCC	1500
Db		1517	GCCCTTTCAAAGCAGAGGCAATGCTTGAGAAATTAAGAGAGCTTGTTGTGTTGGCA	1576
Qy		1501	CCAGGAATTAACATGATCTCGAGTACCAAGTGA	1533

Db 1311 GAGGACTCTCTTGTGGCTGCTCCATCATCTCGATTGGTCCCTAGCTGAACTCAGC 1370
Qy 1321 ACTAGATCCAGTTTAAAGCTGAATGAGGGAAATTCCTCACTCAATCCACCACTGCT 1380
Db 1371 ACTGATATTCAGCTCAAGCCGAGGAGAGGCAATTCATCTGTTCCACCCGGTGGCA 1430
Qy 1381 ACCATTTCTCAGCTATCTGACCAAGGCTCTCTGTTCCACCGGGTACACCACTGGTGAAT 1440
Db 1431 ACCATCTCAGCTACCTAACCAAGCCCTCTTGTACCTCCAGGCACACCGGTCGTCAC 1490
Qy 1441 GCATTGTCAAGACCGCTGCAATGCTGGAATAATATGAGGGCTTGTGTGGATTGGCC 1500
Db 1491 GCACCTGTCAAGCACCGCTGCTGCAATCTTGAAGGGCTGTGTGGATTGGCT 1550
Qy 1501 CCAGAGATTAACATGATCTCGAGTACAAAGTGA 1533
Db 1551 CCAGAAACAAATATGATTTGGAAATACAGTGA 1583

RESULT 11

CPINOIG 1978 bp DNA linear PLN 26-JUN-1995
LOCUS C.paradisi (Macf) INO1 gene.
DEFINITION Z32632.1 GI:602564
ACCESSION Z32632.1 GI:602564
VERSION INO1 gene.
KEYWORDS Citrus x paradisi
SOURCE Citrus x paradisi
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Citrus.

REFERENCE 1 (bases 1 to 1978)
AUTHORS Holland,D.
TITLE Direct Submission
JOURNAL Submitted (08-APR-1994) Doron Holland, Fruit Tree Breeding and
Genetics, Volcani Research, Organization Center, Bet Dagan, 50250,
Israel

FEATURES
source Location/Qualifiers

1..1978
/organism="Citrus x paradisi"
/mol_type="genomic DNA"
/db_xref="taxon:37656"
/clone="C38"
/issue_type="fully expanded leaves"
/dev_stage="juvenile plants"
188..1711
/gene="INO1"
188..1711
/gene="INO1"
/protein_id="CAA83565.1"
/codon_start=1
/db_xref="GI:602565"
/db_xref="GOA:P42802"
/translation="MEIENFKVESPVKXTDHEIHSVVDYETTELVHENRGTQYIV
KERTYVEKTDVHVKLGIVMLVWGNGSDDLITGVLNREGICATDVTYQOANYF
GSLTQASAIRVGSYNGEIEYAPFKSILPMPNDPIVFGGMDISDMNLADAMARAYED
IDLOKQAPRYMSNVPDPIGVDPDPIANOGSRANNVKCTKKKMLQIIXIDIREKE
KAKVDVVVLTANTERYSNVVGLNDYTESLISIDKKEAIESSTLYAICVLENT
PRTNSPONTFVPEGLDLAIRNCLIGSDPKSGQTKMSVLVDPLVGAKIPSTIS
VNHLSNDDNMLISAPQTPRSKEISKSNVDDVWSSNVFPMGLVNRPRMKVPPYAI
BRAMDEYSEIFMGKSTIVLNTCEDSLAAPILIDVLVLAELSTRIGLKEGSGKF
HSFHPVATILSLTKAPLVPRGPVNVNLSKQRAMLENTLRACVGLAPENNNILREK"

ORIGIN

Query Match 68.1%; Score 1044.6; DB 8; Length 1978;
Best Local Similarity 81.9%; Pred. No. 3e-274;
Matches 1255; Conservative 0; Mismatches 269; Indels 9; Gaps 4;

Qy 1 ATGTCATGAGAAATTTAAGTAGAGAGTCCCTAATGTAAGTACCGAGACTGATTT 60
Db 188 ATGTTATGAGAAATTTCAAGGTGTGAGACCCCAATGTGAAGTACAGATCATGAGATT 247

Qy 61 CAGTCGTGTACAACTACGAAACCAACCACTTTGTTACGAGAACAGAAATGGCACTTAT 120
Db 248 CACTCTGTGTACGATTTATGAAACCACTGAGCTTGTTCATGTAGAACAGAAATGGCACTTAT 307
Qy 121 CAGTGAATGTGCAAAACCAATCCGTCAACTACCAATTTAAACCAACCACTATGTCCA 180
Db 308 CAGTGAATGTGCAAAACCACTGCAAAATATGAGTTTCAAAACCACTATGTCCCT 367
Qy 181 AAATGGGGGTGATGCTTGTGGGTGGGTGAAACCAACGGCTCTACCTGACCGGTGT 240
Db 368 AAATGGGGGTGATGCTTGTGGGTGGGTGAAACCAACGGCTCTACCTGACCGGTGT 427
Qy 241 GTTATGCTTAAACAGAGACATTTTATGAGCTACAAAGACAGATTTCAACAGCAAT 300
Db 428 GTTATGCGGAATCGAGAGGAATCTGTGGGCAACAAAGGACATGTGACACCAAGCCAT 487
Qy 301 TACTTTGGCTCCCTACACCAAGCCTGAGCTTATGAGTTGATGCTTCCAGGAGAGGAA 360
Db 488 TATTTGGTTCACTTACCAAGCATACGAAATTAAGTTGGGCTTTCATGAGAGAGAG 547
Qy 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAATGTTTAAATCTGACGACATTTGTGGG 420
Db 548 ATTTATGCTCATTTCAAGAGATTTCTCCATAGGTGAACCAAGATGACATTTGTGTGGA 607
Qy 421 GGAATGGAATATCAGCAATGAACTGAGCTGATGCTCAATGAGGCAAGGCAAGTGTGAC 480
Db 608 GGATGGGACATTTAGTACATGAACTTGTGCTGATGCAATGGCCAGGGCTAGAGTGTAT 667
Qy 481 ATCGATTTGCAAGAGCACTTGAAGGCTTACATGGAATCATGTTTCCACTCCCGGAAATC 540
Db 668 ATTGATCTTCAAAAGCACTTGAAGGCTTACATGGAATCATGTTTCCACTCCCGGAAATC 727
Qy 541 TACGACCCGGAATTCATTGCTGCGCAACCAAGAGAGCGTGCACAAAGTATTAAGGGC 600
Db 728 TATGACCCGGAATTCATTGCTGCGCAATCAAGGGTGGCGTGCACAAAGTATTAAGGGC 787
Qy 601 ACAAGCAAGAGCAAGTTCAGCAATTCATCAAGACATCAAGGCTTTAAGAGCCACC 660
Db 788 ACCAAGAAAGCAAGATCTCAAAATCATCAAGATATTAAGATTAAGAGAGAAAC 847
Qy 661 AAAGTGAACAAGGTGCTGCTGCTGCACTGCAACAGAGAGGTATAGCAATTTGGT 720
Db 848 AAAGTGAACAAGGTGCTGCTGCTGCACTGCAACAGAGGTATAGCAATTTGGT 907
Qy 721 GTAGGCTTAAATGACACCATGAGAAATCTTGGCTGTGAGACAGAAATGAGGCTGAG 780
Db 908 GTGGGGCTAAATGACACCATGAGAAATCTTGGCTGTGAGACAGAAATGAGGCTGAG 967
Qy 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTGATGAGAAATGTTCTTCAATAT 840
Db 968 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTGATGAGAAATGTTCTTCAATAT 1027
Qy 841 GGAACCCCTCAGAAACCTTTTGTACAGAGGCTGATGATCTTGCATGGCGAGAACACT 900
Db 1028 GGAACCCCTCAGAAACCTTTTGTACAGAGGCTGATGATCTTGCATTTAGAGAGAACTGT 1087
Qy 901 TTGATTTGGAGATGACTTCAAGAGTGTGAGACCAAAATGAATCTGTGTGTTGAT 960
Db 1088 TTGATTTGGAGATGACTTCAAGAGTGTGAGACCAAAATGAATCTGTGTGTTGAT 1147
Qy 961 TTTCTTGTGGGGGTGTATACAGGCAATCTATATGTTAGTTTACACCATTTGGGAAC 1020
Db 1148 TTTCTTGTGGGGGTGTATACAGGCAATCTATATGTTAGTTTACACCATTTGGGAAC 1207
Qy 1021 AATGATGTATGAAATCTCGGCTCCAAACCTTCGCTCCAGAGAAATCTCCAGAGC 1080
Db 1208 AATGATGTATGAAATCTCGGCTCCAAACCTTCGCTCCAGAGAAATCTCCAGAGC 1267
Qy 1081 AACGTTGTGACGATATGATGCAACAGATGCAATGCTCTATAGGCTGTGTGAAACATCC 1140
Db 1268 AATGTTGTGATGATGATGCTCCAGCAATG-TATCTTTAAGGCGCTGTGTGAAAC-CCC 1325
Qy 1141 GACCATGTTGTTGTTAATGATGTGCTTACGTAGGGAGATAGCAAGAGACCATGGAT 1200

Db	1326	GACCAACGTTGG-----ATCAAGTAATGCCAATATGTTGG---ATCGAAGAGGCCATGGAT	1378
Qy	1301	GAGTACACTTCAGAGATATTCATGGGTGGAAGAACAACCATTTGTTTGCAACAACATGT	1260
Db	1379	GAGTACACATCAGAAATTTTTCATGGGTGGAAGAGACACCATTTGTCGCAACAACCTGT	1438
Qy	1261	GAGGATTCCTCTTTTAGCTGCTCTAATATCTTGAGCTTGTCCTTCTTGCTAGCTGAGC	1320
Db	1439	GAGGACTCTCTTTTGGCTGCTCCAATCATCTTCAATTTGGTTTGTTGCTGGCTGAACTCAGC	1498
Qy	1321	ACTTGAATCCAGTTTAAAGCTGAAAAATGAGGGAAAATTCACATATTCACCCAGTTGCT	1380
Db	1499	ACCGAGATCCAGCTCAAGCTGGAAGAGAGGGCAAAATTCACCTTTTCAACCTGTGGCT	1558
Qy	1381	ACCATTTCCAGCTATCTGACCAAGGCTCTCTTGTTTCCACCGGGTACCCAGTGGTAAAT	1440
Db	1559	ACAATTTCTCAGCTCACTCACCAGGGCTCTCTGTGTCCTCCAGGCACACCACTGTTGTAAT	1618
Qy	1441	GCAATTGCCAAGCAGCGTGCAAATGCTGAAAAACATAATGAGGGCTTGTGTGAATTTGGCC	1500
Db	1619	GCGCTGTGGAAGCAGCGTGCAAATCTTGGAAGCAATCTGAGGGGCTTGTGTGTGCTTGGCT	1678
Qy	1501	CCAGAGATTAACATGATTTCTCGAGATCAAGTGA	1533
Db	1679	CTTGAGAACCAACATGATTTTGGAAATPACAAGTGA	1711

LOCUS	2053 bp	mRNA	linear	PLN 11-JUL-1996
DEFINITION	Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase			
ACCESSION	U32511			
VERSION	U32511.1			
KEYWORDS	GI:975887			
SOURCE	Mesembryanthemum crystallinum (common iceplant)			
ORGANISM	Mesembryanthemum crystallinum Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Alizocaceae; Mesembryanthemum.			
REFERENCE	1 (bases 1 to 2053)			
AUTHORS	Ishitani,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B., Jensen,R.G. and Bohnert,H.J.			
TITLE	Coordinate transcriptional induction of myo-inositol metabolism during environmental stress			
JOURNAL	Plant J. 9 (4), 537-548 (1996)			
MEDLINE	96208959			
PUBMED	8624516			
REFERENCE	2 (bases 1 to 2053)			
AUTHORS	Ishitani,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B., Jensen,R.G. and Bohnert,H.J.			
TITLE	Direct Submersion			
JOURNAL	Submitted (27-JUL-1995) Hans J. Bohnert, Biochemistry, University of Arizona, Biosciences West 516, Tucson, AZ 85721, USA			
FEATURES	Location/Qualifiers			
source	1..2053			
	/organism="Mesembryanthemum crystallinum"			
	/mol_type="mRNA"			
	/db_xref="taxon:3544"			
	182..1720			
	/codon_start=1			
	/product="myo-inositol-1-phosphate synthase"			
	/protein_id="AAB03687.1"			
	/db_xref="GI:975888"			
	/translation="MFLESPKVESPNVKYTEINIESVNYDTTELVEHNRKQAGYQW IVPKVQYHFKTDTRVPKLGWLVGNGNNGSTLVGVYANREGISWAKDKIQQAN YPSLTQASSIRVGSFNGEETVAPFSLKVPVDPVFGGQVIMDNLADMTARV FDIDLQKLRPMHEHWPLPGIYDPEILANQGSRRNNVTKGKKEQVEVERIDKEF KEKNQDKRVVLVMTGNTERTSNVYQVDMETMLASILEKNESEISPSYALACIE NIFINGSPPMTFPGILIDLAIKNSLIGDPPKSGQTKRKSIVDFPLVGAGIKPSTI VSRNHGNDGNALSLPQTRSKREISNSVNDWVASNGILVPGEGHPDHWVVKYV YVDSDKRADEVTSEIFMGSTNTIVHNTCEDSLAPPIILDLVTLAELSTQIQLAE			
CDS				

EDKDEHSFHPVATILSILTAKPLVPFOTPVVNALSKRAMLENILRACVGLABENMM
LEYK"

Query Match	67.0%	Score 1027.4	DB 8	Length 2053
Best Local Similarity	79.7%	Pred. No. 1.5e-269		
Matches 1227	Conservative 0	Mismatches 306	Indels 6	Gaps 1
Qy	1	ATGTTCAATCCAGAAATTTTAAGTAGAGACATCTTAATATGTAAGTACACCGAGACTGAGATT	60	
Db	182	ATGTTTATTAGAGCTTCAAAGTAGAGGCCCAACGTGAAGTACACAGAAATGAGATT	241	
Qy	61	CAGTCCGTGTAACAATCGAAACCAACCGAATCTGTTACAGAAACAGAA-----TGGC	114	
Db	242	GAATCGGTGTAACAATCGAACACCAACCGAGTGGTTATGAGAAATCGCAAAAGATGCTGCT	301	
Qy	115	ACCATATAGAGGAATGTCATCAACCCCAATCCGTCAATACCAATTTAAACCAACACCAT	174	
Db	302	GGTATATATGATCGTTAAAGCCCAACCGTTCAATATCATCTTCAAGACTGACACTTGT	361	
Qy	175	GTTCCAAAATTTGGGGGTGATGCTTGTGGGTGGGTGAAAACAACGGCTTACCTCAC	234	
Db	362	GTCCCAAACTCGGGGTGATGCTGTGATGGGGTGAATAATACGGTCCACCTGACA	421	
Qy	235	GGTGTGTATTTGTCTAACAGAGAGACATTTATGGGCTACAAAGACATGATTCACAA	294	
Db	422	GGTGTGTGATTCGCCAACCCGAGGGAAATCTTTGGGCAACCAAGGACATGATCAGCA	481	
Qy	295	GCCAAATTAATTTGGGCTCCCTCACCCAAAGCCTCAGCTATTTGAGTTGATCTTCCAGGA	354	
Db	482	GCGAATTAATTTGGTGTGTTGATCCAGGCTTTTCAATTTGGGTCCGGTCTTTTAATGGA	541	
Qy	355	GAGAAATCTATGCCCCATTCAGAGATCTGCTTCAATGTTAATCTGACGACATTTGTG	414	
Db	542	GAGAGATTTATGCTCTCTTCAAGAGCTTCTCCCTATGTTGAACCAAGATGAGTATGTG	601	
Qy	415	TTTGGGGGATGGGATATACGAAACATGAACTCTGTGTATGCCATGGCAGAGGAAAGTGT	474	
Db	602	TTTGGGGGTTGGGACATTAACGACATGAACTGTGTATCAATGACACAGGGCTAGGGTTC	661	
Qy	475	TTTGAATCATGATTTGCGAAGACAGTGGAGGCTTATCAATGATTCATGTTTCCACTCCCC	534	
Db	662	TTTGAATCATGATCTTCAAAAGCAGCTCAGGCTTATCATGAGCAGATGTTTCCCTCCT	721	
Qy	535	GGAATCTACACCCCGATTTCAATTGCTGCGAACCAAGAGAGGCTGCCAACACATGATTT	594	
Db	722	GGTATTTACACCCCTGATTTCAATGCTGCGAACCAAGGCTCCGAGGCAACACATGATTT	781	
Qy	595	AAAGGACCAAAAGACAGAGGTTCCAGCAATCATCAAAACATCAAGGGGTTTAAGAA	654	
Db	782	AAAGGACCAAAAGAGAACAGTTGAGCGATCATTAAGACATTAAGGAGTTCAAGAG	841	
Qy	655	GCCACCAAAAGTGCACAAGTGTGTTGCTCTGTGACTGCCAACACAGAGGTATAGCAAT	714	
Db	842	AAGAACAAAGTGCACAAGGTGTGTGTATGACTGCGAACACAGAGGTACAGCAAT	901	
Qy	715	TTTGTTGTAGGCTTAAATGACACCAATGAGAAATCTTTGGCTGTGTGACACAGAAATGAG	774	
Db	902	GTGTGTGTGGGCTCAACGACACCAATGAGAAATCTGTGGCAATCTCTGGAGAAACAGAG	961	
Qy	775	GCTGAGATTTTCTCTTCAACTGTATGACATTTGACCTGTGTGTATGAGAAATATGTTCCCTTC	834	
Db	962	TGCGAATTTTCTTCAATCCACTTGTATGCTTGGCATGTATCGAAGAAATATTCCTTC	1021	
Qy	835	ATTAAATGAAGCCTCAGAACACTTTTGTATCCAGGGCTGATTTGATCTTGGCATCGGAGG	894	
Db	1022	ATCAATGATGAGCCTCAGAACACTTTTGTGTGCGAGGCTCATGATCTGCTATCAAGAG	1081	
Qy	895	AAACATTTATTTGTGTGAGATGATCTTCAAGAGTGTGCAGACCAAAATGAAATCTGTGTG	954	
Db	1082	AAACATTTATTTGTGTGTATGATCTTCAAGAGTGTGCAGACCAAAATGAAATTTGTCTTC	1141	
Qy	955	GTTGATATTTCTGTGTGGGGCTGTATCAAGCCAAACATCTAATGTTAGTTAACACATCTG	1014	

Db 980 AAGGTAAGTCTCGAAGACCTTTGTCCTCCGTCTCATTTGAATGGCCATTAAGAAAT 1039
Qy 898 ACTTATGATGTTGGAATGATCTTCAAGATGTCAGACCAAAATGAAATCTGTGGTT 957
Db 1040 TGTATTATGGAAGATGATGATCTTCAAGATGTCAGACCTTAAGATGAAATCATGTCGTT 1099
Qy 958 GATTTCTGTGGGGCTGGTATCAAGCCAACTATAGTATGTTACAAACCATCTGGGA 1017
Db 1100 GATTTCTGTGGGGCTGGTATTAAGCCAACTATGTTGATCAACCACTTGGGA 1159
Qy 1018 AACCAATGATGATGATCTCTCGGCTCCAAACCTTCCGCTCAAGAAATCTCCAG 1077
Db 1160 AACCAATGATGATGATCACTTCAAGACCTCAACCTTCAAGATTTCAAAG 1219
Qy 1078 AGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
Db 1220 AGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
Qy 1138 CCGGACCATGTTGTTATTAAGTATGATGATGATGATGATGATGATGATGATGAT 1197
Db 1280 CCGGACCATGTTGTTATTAAGTATGATGATGATGATGATGATGATGATGATGAT 1339
Qy 1198 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
Db 1340 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1399
Qy 1258 TGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
Db 1400 TGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
Qy 1318 AGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377
Db 1460 AGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1519
Qy 1378 GGTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437
Db 1520 GGTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
Qy 1438 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497
Db 1580 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1636
Qy 1498 GGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1531
Db 1637 GGTCCAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1670

RESULT 14
AY028259 1845 bp mRNA linear PLN 26-MAR-2001
LOCUS Avicennia marina myo-inositol 1-phosphate synthase mRNA, complete
DEFINITION cde.
ACCESSION AY028259
VERSION AY028259.1 GI:13447456
KEYWORDS
SOURCE Avicennia marina
ORGANISM Avicennia marina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Lamiales; Acanthaceae; Acanthaceae incertae
sedis; Avicennia.
1 (bases 1 to 1845)
Jithesh,M.N., Parami,M. and Parida,A.
Characterization of a cDNA for myo-inositol 1-phosphate synthase
from the mangrove species Avicennia marina
Unpublished
2 (bases 1 to 1845)
Jithesh,M.N., Parami,M. and Parida,A.
Direct Submission
Submitted (02-MAR-2001) Plant Molecular Biology Lab, M.S.
Swaminathan Research Foundation, III Cross Street, Chennai 600 113,
India

FEATURES
source Location/Qualifiers
1..1845
/organism="Avicennia marina"
/mol_type="mRNA"
/db_xref="taxon:82927"
27..1556
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAK21869.1"
/db_xref="GI:13447457"
/translation="MPISFVKVDSPVKXTDSEIHSVNYETALVELVHNRNGTYQIV
KPTVYERPTDHPVKLGMVLVGMGNNGSTLTGIVIANREGISWATQKQVQAYP
GSLTQASAIRVGSFNGEEIYAPFKSLIPVNPDDIIFGGMDISNMALAMGCGOLD
IDLOQRPRYMHMPPLRGIYYPDPIAAGSRANNVIGTKKEQYQYIKDKERK
KTKVERVVGITANERYSNVVGLNDTASLWASERNAEISPSLTVAIACLEDVP
FLNGSPONTFVGLIELAQRCLIGDDPKSGQTKMSVLVDPLVAGIKPTSIYS
NHGNNDDWMLAPQTFRSKEISKSVVDVMAINSIIEPESHPHVAVIKVPPVG
DKRPMDEVTSIEIFMGKSTIVMHNTECSLLAAPILDLVLAEISTRIOLKABEG
KPHSHFPAVITILSYLTKAPLVPPTPVNVALSKRAMLENILRACVGLAPENNMLEY
K"

ORIGIN
Query Match 66.3%; Score 1016.2; DB 8; Length 1845;
Best Local Similarity 79.4%; Pred. No. 1.7e-266;
Matches 1217; Conservative 0; Mismatches 313; Indels 3; Gaps 1;
Qy 1 ATGTCATCGAAGATTTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
Db 27 ATGTCATCGAAGAGCTTCAAGGTTGACAGCCCAAGTTAAGTACACAGAGTGAATT 86
Qy 61 CAGTCCGTTGATACATCAGAAACCAACCGAATCTTGTTCACAGAAACGAATGACACTAT 120
Db 87 CACTCGTCTACACATGATAAATGCAAGCTTGTTCACAGAAACGAACGGTACTAT 146
Qy 121 CAGTGATGTCAAACCCCAATCCGTCATCAATTTAAACCAACCCCATGTTCCA 180
Db 147 CAGTGATGTCAGAGCCCAACCTGTTAATGAGTTACAGAACCAACCCATGTCGG 206
Qy 181 AAATGGGGGATATGCTTGGGTTGGGGTGAACCAACGGCTTACCTCAACCGGTGT 240
Db 207 AAATTAGGTTATGCTTGGTGGGGAGCAACAAATGTTCAACTGACAGGTGGT 266
Qy 241 GTTATTGCTACAGAGAGACATTTTCATGGGCTACAAAGACAAATTTCAACAGCAAT 300
Db 267 GTTCATTGCAAAATCGGAGGGATTTTCATGGGCAAGAAAGAAAGGTGAACAGGCAAC 326
Qy 301 TACTTGGGCTCCCTCAACCAAGCTCAGCTATTTCAGATGATGATGATGATGATGAT 360
Db 327 TATTTGGGCTCCCTCAACCTCAGGCTTCTGCAATTAAGGTTGGGCTCTTCAATGAGAAAG 386
Qy 361 ATCTATGCCCATTCAGAGATGCTGCTTCCATGTTAATCTGACGACATTTGTTTGG 420
Db 387 ATCTATGCTCCATTCAGAGGCTTACTCCCTATGTCACACCTGATGACATAATTTTGGG 446
Qy 421 GGATGGGATATAGCAACATGAACCTGCTGATGTCACAGGCGCAAGGCTTTTGC 480
Db 447 GGATGGGACATCAGTAATGATGCTGATGTCATGGCATGGCATGGCCCAAGGTTAGAC 506
Qy 481 ATCGATTTGCAGACAGTGAAGGCTTACATGATGATGATGATGATGATGATGATGAT 540
Db 507 ATTTATCTCCAAAGAGCTGAGGCTTACATGAGACATATGTCCTCATCTCCCTGGAAATC 566
Qy 541 TAGACCCGGATTTGATGCTGCGCAACCAAGAGAGCGTGCACCAACGCTGATTAAAGGCG 600
Db 567 TACTACCCGTAATTCATGCTGCGCAATCAGGCTCAGCGGCCAACAAGTATCAAAGGA 626
Qy 601 ACAAGCAAGACCAATTCAGCAATTCATCAAGCATCAAGCGCTTTAAGGAAGCAACC 660
Db 627 ACCAAGAAAGGACCAATTCAGCAATTCATCAAGCATTAAGGAAGTAAAGGAAGAACT 686
Qy 661 AAAGTGAACAAGGTGTTGCTGCTGAGCTGCAACACAGAGAGGTATTAGCAATTTGGTT 720
Db 687 AAGGTGAAGAGGTAGTG--TGGGAGACAGCCCAACCGAGATACAGCAATGTTT 743

QY	721	GTAGGCTTAATGACACCATGGAATCTCTGGCTGTGGAAGAGAGCTGAG	780
DB	744	GTGGGCTTAATGATACAGCGAAGCCCTTAATGCTCTGTAAGAAATAGAGCTGAG	803
QY	781	ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAAAATGTTCTTTCAATAT	840
DB	804	ATATCTCTTCCACCTTAATATGCTAGCTGTGTGATGAAAATGTTCTTTCAATAT	863
QY	841	GGAAGCCCTCAGAACCTTTGTACAGAGGCTGATGATCTTGCCATCCGAGGAACAAT	900
DB	864	GGAAGCCCTCAGAACCTTTGTACAGAGGCTGATGATCTTGCCATCCGAGGAACAAT	923
QY	901	TGATTTGTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGATGAT	960
DB	924	TGATTTGTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGATGAT	983
QY	961	TTTCTTGTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGATGAT	1020
DB	984	TTTCTTGTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGATGAT	1043
QY	1021	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1080
DB	1044	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1103
QY	1081	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1140
DB	1104	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1163
QY	1141	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1200
DB	1164	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1223
QY	1201	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1260
DB	1224	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1283
QY	1261	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1320
DB	1284	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1343
QY	1321	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1380
DB	1344	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1403
QY	1381	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1440
DB	1404	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1463
QY	1441	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1500
DB	1464	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1523
QY	1501	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1533
DB	1524	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1556

RESULT 15
 AX054630 1781 bp DNA 11near PAT 13-JAN-2001
 LOCUS AX054630
 DEFINITION Sequence 1 from Patent WO0073473.
 ACCESSION AX054630
 VERSION AX054630.1 GI:12228190
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica napus (rape)
 Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE
 AUTHORS Georges, F., Hussain, A.A. and Keller, W.A.
 TITLES Method for reducing phytate in canola meal using genetic

JOURNAL		manipulation involving myo-inositol 1-phosphate synthase gene	
FEATURES		Patent: WO 0073473-A 1 07-DEC-2000;	
source		NATIONAL RESEARCH COUNCIL CANADA (CA)	
1. 1781		Location/Qualifiers	
/organism="Brassica napus"			
/mol_type="unassigned DNA"			
/db_xref="taxon:3708"			
ORIGIN			
Query Match		65.2%; Score 1000.2; DB 6; Length 1781;	
Best Local Similarity		78.3%; Pred. No. 4.2e-262;	
Matches 1200; Conservative		0; Mismatches 333; Indels 0; Gaps 0;	
QY	1	ATGTCATCGAGAAATTTAAGGTAGAGTCTCTAATGAAATGACCGAGATGATG	60
DB	57	ATGTCATCGAGAGCTTCAAAGTCCAGAGCCCGAAGCGTAAATACCGAGATGAGTT	116
QY	61	CAGTCGGGTACAACTACGAAACACCGAATCTGTTCAAGAGACGAAATGCACTAT	120
DB	117	CATTCGGGTACGATTTACGAGACACGAGAGTCTTCAAGAGACGTAACGCTGTTAC	176
QY	121	CAGTGAATGTCAAACCCAAATCCGTCAATACCAATTTAAACCAACCAATTTTCA	180
DB	177	CAGTGAATGTCGAAACCCAAAGTTGTCAAAATACGATTTCAAAACCGAACCTGTCCTC	236
QY	181	AAATGGGGGTGATGCTTGTGGGTGGGTGGAACCAACGCTTACCTCACCGGTGT	240
DB	237	AAATGGGGGTGATGCTTGTGGGTGGGTGGAACCAACGCTTACCTCACCGGTGT	296
QY	241	GTTATGCTTACAGAGAGACATTTTATGGGTACAAAGAGACAAATTCACAGCCAT	300
DB	297	GTTATGCTTACAGAGAGACATTTTATGGGTACAAAGAGACAAATTCACAGCCAT	356
QY	301	TACTTGGGTCCCTACACCAAGCTCAGCTATTCAGATGATGATCCTTCCAGGAGAGA	360
DB	357	TACTTGGGTCCCTACACCAAGCTCAGCTATTCAGATGATGATCCTTCCAGGAGAGA	416
QY	361	ATCTAAGCCCATTCAGAGTCTGCTCCAAATGTTAAATCTGACAGCATTTGTGAGG	420
DB	417	ATCTAAGCCCATTCAGAGTCTGCTCCAAATGTTAAATCTGACAGCATTTGTGAGG	476
QY	421	GGATGGGATATCAGCAACATGAACTGGTGTGATGCTGATGCTGAGGCAAGGTTTGA	480
DB	477	GGATGGGATATCAGCAACATGAACTGGTGTGATGCTGATGCTGAGGCAAGGTTTGA	536
QY	481	ATCGATTTGCAAGACAGTGAAGGCTTACATGAAATCCATGCTTCCATCCCGGAAATC	540
DB	537	ATCGATTTGCAAGACAGTGAAGGCTTACATGAAATCCATGCTTCCATCCCGGAAATC	596
QY	541	TACGACCGGATTTTATGCTGCTCCAAACAGAGAGCGTGCACAAACGTTGATTAAGGC	600
DB	597	TACGACCGGATTTTATGCTGCTCCAAACAGAGAGCGTGCACAAACGTTGATTAAGGC	656
QY	601	ACAAAGCAAGCAAGTTCAGCAATCATCAAGACATCAAGGCTTTAAGAGCCACC	660
DB	657	ACAAAGCAAGCAAGTTCAGCAATCATCAAGACATCAAGGCTTTAAGAGCCACC	716
QY	661	AAAGTGAAGCAAGTGTGCTCTGTGACGCTGCAACAGAGAGTATGCAATTTGGTT	720
DB	717	AAAGTGAAGCAAGTGTGCTCTGTGACGCTGCAACAGAGAGTATGCAATTTGGTT	776
QY	721	GTAGGCTTAATGACACCATGAGAGATCTCTGGCTGTGATGAGAGAAATGAGGCTGAG	780
DB	777	GTAGGCTTAATGACACCATGAGAGATCTCTGGCTGTGATGAGAGAAATGAGGCTGAG	836
QY	781	ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATAT	840
DB	837	ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATAT	896
QY	841	GGAAGCCCTCAGAACCTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAACACT	900
DB	897	GGAAGCCCTCAGAACCTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAACACT	956

```
OY 901 TTGATTGGATGAGATGACTTCAAGAGTGATCAGACCAGAAATGAATCTGTGTTGGTAT 960
DB 957 TTGATCGGTGGAGATGACTTCAAGAGTGATCAGACCAGAAATGAATCTGTGTTGGTAT 1016
OY 961 TTTCTTGTGGGGCTGGTATCAAGCCAACTCTATAGTTAGTTACAACCATCTGGGAAC 1020
DB 1017 TTCTTGTGTGTCAGGCAATCAAGCCCTACTCAATTGTGAGCTACATCACTAGGGAAC 1076
OY 1021 AATGATGATGATGATCTCTGCGCTCCAAACCTTCGCTCCAAAGAAATCTCCAAAGAC 1080
DB 1077 AACGATGAGATGAACCTTCAGCTCCACAGACATTCAGATCTAAGGAGATCTCAAAAGT 1136
OY 1081 AACGTTGTGACGATGATGATCAACAGCAATGCCATCTCTATAGAGCTGGTGAACATCCC 1140
DB 1137 AATGTGTTGACATATGTTGCTAGCAGATCTCTTCGAGCCCGGGGAACATCCA 1196
OY 1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGATAGCAAGAGCCATGAT 1200
DB 1197 GACCATGTTGTTGTTATCAAGTATGTACGTATGTTGCAGATAGCAAGAGCCATGAT 1256
OY 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAACACCATGTTTGCAACAACATGT 1260
DB 1257 GAGTATACATCAGAGATATTCATGGGAGGCAAGAACACATGTGATGCACAATACCTGC 1316
OY 1261 GAGGATTCCTTTTACCTGCTCCTATTTATCTTGGACTTGTCTTGTGAGCTGAGC 1320
DB 1317 GAGGACTCTCTTACCTGCTCCTCAATCTTGGATCTTGTCTCTCGTGAATCAGC 1376
OY 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGGAAATTCACATTCACCCAGTTGCT 1380
DB 1377 ACCAGATTCAGTTCAAAATCCAGAAAGAGGGAAAGTTTCAATCTTCCATCTGTGGCC 1436
OY 1381 ACCATTCAGCTATCTGACCAAGGCTCTCTGTTCACCGGGTACACCAAGTGTGAAT 1440
DB 1437 ACCAACTTAGCTATCTCACCAAGGCAACGCTGTGCCCGGGGAACACCGGTGTTAAT 1496
OY 1441 GCATTGTCAAGAGGAGGCAATGCTGGAACATTAAGAGGCTGTGTGATTTGGCC 1500
DB 1497 GCGCTGTCCAGAGGAGGAGGCTATGCTGAGAAATATTTCTTAGGGCGGTGTGGCG 1556
OY 1501 CCAGAGAAATACATGATTCGAGATCAAGTGA 1533
DB 1557 CCGGAGAACATATGATCTTGGAATCAAGTGA 1589
```

Search completed: June 8, 2005, 02:54:03
Job time : 6693.92 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:46:53 ; Search time 81.5 Seconds
(without alignments)
2254.032 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MFIEHFKECPNVKXTEETEI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1599520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2632	100.0	510	US-10-025-003-2	Sequence 2, Appli
2	2632	100.0	510	US-10-025-003-10	Sequence 10, Appli
3	2632	100.0	510	US-10-718-952-2	Sequence 2, Appli
4	2632	100.0	510	US-10-718-952-10	Sequence 10, Appli
5	2627	99.8	510	US-10-025-003-6	Sequence 6, Appli
6	2627	99.8	510	US-10-718-952-6	Sequence 6, Appli
7	2621.5	99.5	511	US-10-424-599-213009	Sequence 213009,
8	2600	98.8	510	US-10-025-003-14	Sequence 14, Appli
9	2600	98.8	510	US-10-025-003-16	Sequence 16, Appli
10	2600	98.8	510	US-10-718-952-14	Sequence 14, Appli
11	2600	98.8	510	US-10-718-952-16	Sequence 16, Appli
12	2593	98.5	510	US-10-025-003-12	Sequence 12, Appli

13	2593	98.5	510	US-10-718-952-12	Sequence 12, Appli
14	2524	95.9	510	US-10-424-599-154864	Sequence 154864, A
15	2432	92.4	505	US-10-442-017-15	Sequence 15, Appli
16	2356	89.5	510	US-10-425-115-231852	Sequence 231852, A
17	2356	89.5	516	US-10-425-114-62568	Sequence 62568, A
18	2355	89.5	510	US-10-767-701-46278	Sequence 46278, A
19	2354	89.4	510	US-10-739-930-7635	Sequence 7635, Ap
20	2351	89.3	510	US-10-425-115-306814	Sequence 306814, A
21	2347	89.2	510	US-10-437-963-173252	Sequence 173252, A
22	2345	89.1	542	US-10-425-114-58674	Sequence 58674, A
23	2336	88.8	510	US-09-921-232-11	Sequence 11, Appli
24	2336	88.8	510	US-09-921-330-11	Sequence 11, Appli
25	2336	88.8	510	US-09-921-329-11	Sequence 11, Appli
26	2335	88.7	510	US-09-727-628-2	Sequence 2, Appli
27	2335	88.7	510	US-10-425-115-231850	Sequence 231850, A
28	2335	88.7	510	US-10-425-115-231857	Sequence 231857, A
29	2335	88.7	512	US-10-425-114-39621	Sequence 39621, A
30	2335	88.7	512	US-10-425-114-53231	Sequence 53231, A
31	2335	88.7	512	US-10-425-114-66216	Sequence 66216, A
32	2334	88.7	510	US-10-425-115-231853	Sequence 231853, A
33	2334	88.7	512	US-10-425-114-46916	Sequence 46916, A
34	2205.5	83.8	509	US-10-437-963-187768	Sequence 187768, A
35	2193	83.3	510	US-10-425-115-268216	Sequence 268216, A
36	2193	83.3	521	US-10-425-114-62889	Sequence 62889, A
37	2193	83.3	522	US-10-425-114-62888	Sequence 62888, A
38	2193	83.3	529	US-10-425-114-58895	Sequence 58895, A
39	2129	80.9	458	US-10-425-115-231843	Sequence 231843, A
40	2104	79.9	450	US-10-425-115-231845	Sequence 231845, A
41	2016	76.6	431	US-10-424-599-154863	Sequence 154863, A
42	1607	61.1	558	US-09-876-889-348	Sequence 348, Appli
43	1596.5	60.7	534	US-10-128-714-3515	Sequence 3515, Ap
44	1592.5	60.5	534	US-10-128-714-8515	Sequence 8515, Ap
45	1558	59.2	332	US-10-425-114-42702	Sequence 42702, A

ALIGNMENTS

RESULT 1
US-10-025-003-2
; Sequence 2, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-2

Query Match 100.0%; Score 2632; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 8.6e-220; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFIEHFKECPNVKXTEETEIQSVNYETTELVDHNRNGTQWIVPKSVKYEFTKTHHP 60
DB 1 MFIEHFKECPNVKXTEETEIQSVNYETTELVDHNRNGTQWIVPKSVKYEFTKTHHP 60
QY 61 KLGVLTWVGNGNSTLTGVIARREGISMATKDKIQQANTFGSLTQSAIRVGSFQEE 120

Db	61	KLGWLVGCGNNGSTLTGCVIANEGISMAKXKQIQANYPGSLTQASAIRVOSFGEE	120
Qy	121	IYAPFKSLIPWVNPDDIVFGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLGI	180
Db	121	IYAPFKSLIPWVNPDDIVFGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLGI	180
Qy	181	YBDFPFAAQQEERANNVIGTKQEOVOQIQIKAKFKATVVDKVVVLTANTRYSNLV	240
Db	181	YBDFPFAAQQEERANNVIGTKQEOVOQIQIKAKFKATVVDKVVVLTANTRYSNLV	240
Qy	241	VGANDMENLAAVDNEAEISPSTLYAACMEVNPPTNSPONTVPGLIDATARNT	300
Db	241	VGANDMENLAAVDNEAEISPSTLYAACMEVNPPTNSPONTVPGLIDATARNT	300
Qy	301	LIGGDFKSGQTKMSVLVDFVCGAIGKPTISVSNHGNNDGNMLSNAPQTFKSEIKS	360
Db	301	LIGGDFKSGQTKMSVLVDFVCGAIGKPTISVSNHGNNDGNMLSNAPQTFKSEIKS	360
Qy	361	NYVDDMVNSNALYEBGEHPDHVVYIKVPYVYGSKRAMDEYTSEIFMGKSTIVLHNTC	420
Db	361	NYVDDMVNSNALYEBGEHPDHVVYIKVPYVYGSKRAMDEYTSEIFMGKSTIVLHNTC	420
Qy	421	EDSLAAPITLIDVLAELESTRIEKFAENEGKFFSHFVAITLISYLKAPLVPFGTVPVN	480
Db	421	EDSLAAPITLIDVLAELESTRIEKFAENEGKFFSHFVAITLISYLKAPLVPFGTVPVN	480
Qy	481	ALSKORAMLENIRACVGLAPENNMILEYK 510	
Db	481	ALSKORAMLENIRACVGLAPENNMILEYK 510	

```

RESULT 2
US-10-025-003-10
; Sequence 10, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streifel, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-10

```

Query Match	100.0%;	Score 2632;	DB 14;	Length 510;
Best Local Similarity	100.0%;	Pred. No. 8.6e-220;		
Matches 510; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	Db	Qy	Db
1 MFLENFVKEPNKYETETIQSVYNNETTELVAHENRNTGYQIMPKPSVKKEFFNTNHP	1 MFLENFVKEPNKYETETIQSVYNNETTELVAHENRNTGYQIMPKPSVKKEFFNTNHP	121 IYAPKPSLLPYNVPDDIVFEGMDISNMNLADAMARAKYFDIDLQKQRPVYESMLPLPGI	121 IYAPKPSLLPYNVPDDIVFEGMDISNMNLADAMARAKYFDIDLQKQRPVYESMLPLPGI
60	60	180	180
61 KLGVMVLGVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGQEE	61 KLGVMVLGVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGQEE	121 IYAPKPSLLPYNVPDDIVFEGMDISNMNLADAMARAKYFDIDLQKQRPVYESMLPLPGI	121 IYAPKPSLLPYNVPDDIVFEGMDISNMNLADAMARAKYFDIDLQKQRPVYESMLPLPGI
120	120	180	180
62 KLGVMVLGVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGQEE	62 KLGVMVLGVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGQEE	121 IYAPKPSLLPYNVPDDIVFEGMDISNMNLADAMARAKYFDIDLQKQRPVYESMLPLPGI	121 IYAPKPSLLPYNVPDDIVFEGMDISNMNLADAMARAKYFDIDLQKQRPVYESMLPLPGI
120	120	180	180

Qy	181	YBDDFLAANOERANNVJKTKOEVOOIIIMDIAFKKATVDDCVVYLWMTANERSNLV	240
Db	181	YDDFLAANOERANNVJKTKOEVOOIIIMDIAFKKATVDDCVVYLWMTANERSNLV	240
Qy	241	VGANDTMENILAAVDRNEAETSPSTLYAIACMENVPFINGSPONTFVPGILDAIARNT	300
Db	241	VGANDTMENILAAVDRNEAETSPSTLYAIACMENVPFINGSPONTFVPGILDAIARNT	300
Qy	301	LIGGDDFKSGQTKMSVLVDFLVGAGIKPTSIIVSYNHLGNNDGNLSAPOTFSKEISKS	360
Db	301	LIGGDDFKSGQTKMSVLVDFLVGAGIKPTSIIVSYNHLGNNDGNLSAPOTFSKEISKS	360
Qy	361	NVDDMVNSNALIYBEGEHPDHVVITKVPYVGSKRAMDEYTSIEMGKSTIVLHNTC	420
Db	361	NVDDMVNSNALIYBEGEHPDHVVITKVPYVGSKRAMDEYTSIEMGKSTIVLHNTC	420
Qy	421	EDSLAALPIITLDVLVLAELSTRIEFKARNECKFPHFVATITLSYLTKALVPPGPVVN	480
Db	421	EDSLAALPIITLDVLVLAELSTRIEFKARNECKFPHFVATITLSYLTKALVPPGPVVN	480
Qy	481	ALSKORAMLENTIRACVGLAPENNITLEYK	510
Db	481	ALSKORAMLENTIRACVGLAPENNITLEYK	510

```

RESULT 3
US-10-718-952-2
; Sequence 2, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-2

```

Query Match	100.0%	Score 2632;	DB 16;	length 510;
Best Local Similarly	100.0%	Pred. NO. 8, 6e-220;		
Matches 510; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MEINFEKCECNVAKYETEIEIOSVYNVETETELVHNRNGTYOMIYKPSVKVEFKTNIHVP	60
QY	1	MEINFEKCECNVAKYETEIEIOSVYNVETETELVHNRNGTYOMIYKPSVKVEFKTNIHVP	60
Db	1	MFIEINFKECCNVAKYETEIEIOSVYNVETETELVHNRNGTYOMIYKPSVKVEFKTNIHVP	60
QY	61	KLGVMLVGMGNNCSSTLTGIVIANREGISWATKDKIOOANYFGSLTOASAIRVGSFOGEE	120
Db	61	KLGVMLVGMGNNCSSTLTGIVIANREGISWATKDKIOOANYFGSLTOASAIRVGSFOGEE	120
QY	121	IYAPFKSLLPVNVNDDIVFGSGMDSNNMLADAMARAKVFDIDLKOLRPVMSMLPLPGI	180
Db	121	IYAPFKSLLPVNVNDDIVFGSGMDSNNMLADAMARAKVFDIDLKOLRPVMSMLPLPGI	180
QY	181	YDPDFEIANOGERRANNVIKGTFOEQOVQOIIKDIAFKFATETVDSRVVLMNTANTERSNLT	240
Db	181	YDPDFEIANOGERRANNVIKGTFOEQOVQOIIKDIAFKFATETVDSRVVLMNTANTERSNLT	240
QY	241	VGUNDYMENTLAAVDRNEAEISPSLTAYAIACWENVPFINGSPONTFVGLIDLAIRANT	300


```

Db 241 VGLNDTMENMLAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLDVFLVAGAIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLDVFLVAGAIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Qy 421 EDSLAAPIILDLVLLAESTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480
Db 421 EDSLAAPIILDLVLLAESTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480
Qy 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

```

RESULT 4

```

US-10-718-952-10
; Sequence 10, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelt, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRF
; ORGANISM: Glycine max
US-10-718-952-10

```

Query Match 100.0%; Score 2632; DB 16; Length 510;
 Best Local Similarity 100.0%; Pred. No. 8.6e-220;
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MFIENFKVECPNVKKTETETIQSVNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Db 1 MFIENFKVECPNVKKTETETIQSVNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Qy 61 KLGWLVGVGNGNSTLTGVTANREGISWATKDKIQOANYFGSLTQASAIRVGSQGE 120
Db 61 KLGWLVGVGNGNSTLTGVTANREGISWATKDKIQOANYFGSLTQASAIRVGSQGE 120
Qy 121 IYAPFKSLPVMNPPDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPFYMSMLPLPGI 180
Db 121 IYAPFKSLPVMNPPDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPFYMSMLPLPGI 180
Qy 181 YDPDFIAANQERANNVIKGTQEOVOQIIKOIKAFKATKVDKVVVLTANTERTYSLV 240
Db 181 YDPDFIAANQERANNVIKGTQEOVOQIIKOIKAFKATKVDKVVVLTANTERTYSLV 240
Qy 241 VGLNDTMENMLAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNT 300
Db 241 VGLNDTMENMLAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLDVFLVAGAIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLDVFLVAGAIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360

```

```

Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Qy 421 EDSLAAPIILDLVLLAESTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480
Db 421 EDSLAAPIILDLVLLAESTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480
Qy 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

```

RESULT 5

```

US-10-025-003-6
; Sequence 6, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelt, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRF
; ORGANISM: Glycine max
US-10-025-003-6

```

Query Match 99.8%; Score 2627; DB 14; Length 510;
 Best Local Similarity 99.8%; Pred. No. 2.3e-219;
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 MFIENFKVECPNVKKTETETIQSVNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Db 1 MFIENFKVECPNVKKTETETIQSVNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Qy 61 KLGWLVGVGNGNSTLTGVTANREGISWATKDKIQOANYFGSLTQASAIRVGSQGE 120
Db 61 KLGWLVGVGNGNSTLTGVTANREGISWATKDKIQOANYFGSLTQASAIRVGSQGE 120
Qy 121 IYAPFKSLPVMNPPDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPFYMSMLPLPGI 180
Db 121 IYAPFKSLPVMNPPDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPFYMSMLPLPGI 180
Qy 181 YDPDFIAANQERANNVIKGTQEOVOQIIKOIKAFKATKVDKVVVLTANTERTYSLV 240
Db 181 YDPDFIAANQERANNVIKGTQEOVOQIIKOIKAFKATKVDKVVVLTANTERTYSLV 240
Qy 241 VGLNDTMENMLAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNT 300
Db 241 VGLNDTMENMLAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLDVFLVAGAIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLDVFLVAGAIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Qy 421 EDSLAAPIILDLVLLAESTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTTPVNN 480

```

Db 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSHFHPATILSYLTAKALVPFGTPTVVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510

RESULT 6
US-10-718-952-6

/ Sequence 6, Application US/10718952
/ Publication No. US20040128713A1
/ GENERAL INFORMATION:
/ APPLICANT: Hitz, William
/ APPLICANT: Sebastian, Scott
/ APPLICANT: Grace, John
/ APPLICANT: Streitz, Leon
/ TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
/ FILE REFERENCE: BB-1077-C
/ CURRENT APPLICATION NUMBER: US/10/718,952
/ PRIOR FILING DATE: 2003-11-21
/ PRIOR APPLICATION NUMBER: 08/835,751
/ PRIOR FILING DATE: APRIL 8, 1997
/ PRIOR APPLICATION NUMBER: PCT/US98/06822
/ PRIOR FILING DATE: APRIL 7, 1998
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 6
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Glycine max
US-10-718-952-6

Query Match 99.8%; Score 2627; DB 16; Length 510;
Best Local Similarity 99.8%; Pred. No. 2,3e-219;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MFIENFKCECPNVKKTETETESVYNYVETTELVHNRNGTYQWIVKPKSVKTEFKTNHVP 60
Db 1 MFIENFKCECPNVKKTETETESVYNYVETTELVHNRNGTYQWIVKPKSVKTEFKTNHVP 60
Qy 61 KLGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
Db 61 KLGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
Qy 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKVPDIDLOKQLRPYMESMLPLPG 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKVPDIDLOKQLRPYMESMLPLPG 180
Qy 181 YDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKEATKVDKVVVLTANTERYSNLV 240
Db 181 YDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKEATKVDKVVVLTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACVMENVPFINGSPONTFVPGILDLAIARN 300
Db 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACVMENVPFINGSPONTFVPGILDLAIARN 300
Qy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGNLSAPQTFRSKEISK 360
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGNLSAPQTFRSKEISK 360
Qy 361 NVVDVMNNSNALIYEPGEHPDHVVVVKYVPYVGSKRAMEDEYSEIFMGKSTIVLHNT 420
Db 361 NVVDVMNNSNALIYEPGEHPDHVVVVKYVPYVGSKRAMEDEYSEIFMGKSTIVLHNT 420
Qy 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSHFHPATILSYLTAKALVPFGTPTVV 480
Db 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSHFHPATILSYLTAKALVPFGTPTVV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510

RESULT 7

US-10-424-599-213009

/ Sequence 213009, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 213009
/ LENGTH: 511
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(511)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.pcp
US-10-424-599-213009

Query Match 99.6%; Score 2621.5; DB 15; Length 511;
Best Local Similarity 99.8%; Pred. No. 7e-219;
Matches 510; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MFIENFKCECPNVKKTETETESVYNYVETTELVHNRNGTYQWIVKPKSVKTEFKTNHVP 60
Db 1 MFIENFKCECPNVKKTETETESVYNYVETTELVHNRNGTYQWIVKPKSVKTEFKTNHVP 60
Qy 61 KLGVMLVGMG-GNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGE 119
Db 61 KLGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
Qy 120 EYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKVPDIDLOKQLRPYMESMLPLPG 179
Db 120 EYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKVPDIDLOKQLRPYMESMLPLPG 180
Qy 180 IYDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKEATKVDKVVVLTANTERYSNLV 239
Db 180 IYDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKEATKVDKVVVLTANTERYSNLV 240
Qy 240 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACVMENVPFINGSPONTFVPGILDLAIARN 299
Db 240 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACVMENVPFINGSPONTFVPGILDLAIARN 300
Qy 300 TLIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGNLSAPQTFRSKEISK 359
Db 300 TLIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGNLSAPQTFRSKEISK 360
Qy 360 SNVDVMNNSNALIYEPGEHPDHVVVVKYVPYVGSKRAMEDEYSEIFMGKSTIVLHNT 419
Db 360 SNVDVMNNSNALIYEPGEHPDHVVVVKYVPYVGSKRAMEDEYSEIFMGKSTIVLHNT 420
Qy 420 CEDSLAAPIILDLVLAELSTRIEFKAENEGKFSHFHPATILSYLTAKALVPFGTPTVV 479
Db 420 CEDSLAAPIILDLVLAELSTRIEFKAENEGKFSHFHPATILSYLTAKALVPFGTPTVV 480
Qy 480 NALSKORAMLENIMRACVGLAPENNMIIEYK 510
Db 480 NALSKORAMLENIMRACVGLAPENNMIIEYK 511

RESULT 8
US-10-025-003-14
/ Sequence 14, Application US/10025003

```

; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-14

Query Match          98.8%; Score 2600; DB 14; Length 510;
Best Local Similarity 98.6%; Pred. No. 5.2e-217;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPNVKTETETIQSVNYVETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
DB 1 MEIENFKVCEPNVKTETETIQSVNYVETTELVEHNRNGTYQWIVKPKSVVYQFKTHTVP 60
QY 61 KIGWMLVGWGNNGSTLTGVIANREGISMATKDKIQOANYFGSLTQASAIRVGSFGSBE 120
DB 61 KIGWMLVGWGNNGSTLTGVIANREGISMATKDKIQOANYFGSLTQASAIRVGSFGSBE 120
QY 121 IYAPFKSLPMTNPDIVFGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
DB 121 IYAPFKSLPMTNPDIVFGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
QY 181 YPDPFLAANQEBRANNVIKGTQEQVOQIIKDIAKFEATKYDKVVLWTANTERYSNLV 240
DB 181 YPDPFLAANQEBRANNVIKGTQEQVOQIIKDIAKFEATKYDKVVLWTANTERYSNLV 240
QY 241 VGLNTMNTMULLAAYDRNEAISPSTLYAIAACWENVPFNGSPONTFVGLIDLAIRNT 300
DB 241 VGLNTMNTMULLAAYDRNEAISPSTLYAIAACWENVPFNGSPONTFVGLIDLAIRNT 300
QY 301 LIGGDDFKSGQTKMKSVDVFLVAGIKPTSIYSYHNLGNNDGMNLSAPQTRSKSISKS 360
DB 301 LIGGDDFKSGQTKMKSVDVFLVAGIKPTSIYSYHNLGNNDGMNLSAPQTRSKSISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPITLDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKPLVPPGTPPVN 480
DB 421 EDSLLAAPITLDVLVLAELSTRIOFKAENEGKFSHPVATILSYLTAKPLVPPGTPPVN 480
QY 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

RESULT 9
US-10-025-003-16
; Sequence 16, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon

```

```

; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-16

Query Match          98.8%; Score 2600; DB 14; Length 510;
Best Local Similarity 98.6%; Pred. No. 5.2e-217;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPNVKTETETIQSVNYVETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
DB 1 MEIENFKVCEPNVKTETETIQSVNYVETTELVEHNRNGTYQWIVKPKSVVYQFKTHTVP 60
QY 61 KIGWMLVGWGNNGSTLTGVIANREGISMATKDKIQOANYFGSLTQASAIRVGSFGSBE 120
DB 61 KIGWMLVGWGNNGSTLTGVIANREGISMATKDKIQOANYFGSLTQASAIRVGSFGSBE 120
QY 121 IYAPFKSLPMTNPDIVFGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
DB 121 IYAPFKSLPMTNPDIVFGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
QY 181 YPDPFLAANQEBRANNVIKGTQEQVOQIIKDIAKFEATKYDKVVLWTANTERYSNLV 240
DB 181 YPDPFLAANQEBRANNVIKGTQEQVOQIIKDIAKFEATKYDKVVLWTANTERYSNLV 240
QY 241 VGLNTMNTMULLAAYDRNEAISPSTLYAIAACWENVPFNGSPONTFVGLIDLAIRNT 300
DB 241 VGLNTMNTMULLAAYDRNEAISPSTLYAIAACWENVPFNGSPONTFVGLIDLAIRNT 300
QY 301 LIGGDDFKSGQTKMKSVDVFLVAGIKPTSIYSYHNLGNNDGMNLSAPQTRSKSISKS 360
DB 301 LIGGDDFKSGQTKMKSVDVFLVAGIKPTSIYSYHNLGNNDGMNLSAPQTRSKSISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPITLDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKPLVPPGTPPVN 480
DB 421 EDSLLAAPITLDVLVLAELSTRIOFKAENEGKFSHPVATILSYLTAKPLVPPGTPPVN 480
QY 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

RESULT 10
US-10-718-952-14
; Sequence 16, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21

```

;; PRIOR FILING DATE: APRIL 8, 1997
;; PRIOR APPLICATION NUMBER: PCT/US98/06822
;; PRIOR FILING DATE: APRIL 7, 1998
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 14
;; LENGTH: 510
;; TYPE: PRT
;; ORGANISM: Glycine max
US-10-718-952-14

Query Match 98.8%; Score 2600; DB 16; Length 510;
Best Local Similarity 98.6%; Pred. No. 5.2e-217;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MFIENFKVCEPNVKTETETIOSVYNYVETTELVEHNRNGTYOMIVKPKSVKYEFTKNTIHVP 60
DB 1 MFIENFKVCEPNVKTETETIOSVYNYVETTELVEHNRNGTYOMIVKPKSVNYQFKNTIHVP 60
QY 61 KLGWLVGMGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGBE 120
DB 61 KLGWLVGMGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGBE 120
QY 121 IYAPFKSLIPVNPDPDIYFGGWDISNNMLADAMAAKVPDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPVNPDPDIYFGGWDISNNMLADAMAAKVPDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEERANNVIGTKOEVOQIIDKIKAFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPDFIAANOEERANNVIGTKOEVOQIIDKIKAFKATKVDKVVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDRNEAEISPTLYAIACWENVPIFGSPONTFVPGILDIAIANT 300
DB 241 VGLNDTMENLLAAVDRNEAEISPTLYAIACWENVPIFGSPONTFVPGILDIAIANT 300
QY 301 LIGGDDFKSGQTKMKSVLVDLVGAGIKPTISVYNHGNNDGMLSAPOFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDLVGAGIKPTISVYNHGNNDGMLSAPOFRSKEISKS 360
QY 361 NVDDMVNSNALIYEPGEHPDHVVYIKVPYVGSKRAMDEYSEIFMGKSTIYLAHTC 420
DB 361 NVDDMVNSNALIYEPGEHPDHVVYIKVPYVGSKRAMDEYSEIFMGKSTIYLAHTC 420
QY 421 EDSLAAPIIIDLVLLAEIETRIEFKAENEGKFSHPVATILSYLTAPLVPPTPVVN 480
DB 421 EDSLAAPIIIDLVLLAEIETRIEFKAENEGKFSHPVATILSYLTAPLVPPTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 11
US-10-718-952-16
;; Sequence 16, Application US/10718952
;; Publication No. US20040128713A1
;; GENERAL INFORMATION:
;; APPLICANT: Hitz, William
;; APPLICANT: Sebastian, Scott
;; APPLICANT: Grace, John
;; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
;; FILE REFERENCE: BB-1077-C
;; CURRENT APPLICATION NUMBER: US/10/718,952
;; PRIOR FILING DATE: 2003-11-21
;; PRIOR APPLICATION NUMBER: 08/835,751
;; PRIOR FILING DATE: APRIL 8, 1997
;; PRIOR APPLICATION NUMBER: PCT/US98/06822
;; PRIOR FILING DATE: APRIL 7, 1998
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 16

;; LENGTH: 510
;; TYPE: PRT
;; ORGANISM: Glycine max
US-10-718-952-16

Query Match 98.8%; Score 2600; DB 16; Length 510;
Best Local Similarity 98.6%; Pred. No. 5.2e-217;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MFIENFKVCEPNVKTETETIOSVYNYVETTELVEHNRNGTYOMIVKPKSVKYEFTKNTIHVP 60
DB 1 MFIENFKVCEPNVKTETETIOSVYNYVETTELVEHNRNGTYOMIVKPKSVNYQFKNTIHVP 60
QY 61 KLGWLVGMGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGBE 120
DB 61 KLGWLVGMGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGBE 120
QY 121 IYAPFKSLIPVNPDPDIYFGGWDISNNMLADAMAAKVPDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPVNPDPDIYFGGWDISNNMLADAMAAKVPDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEERANNVIGTKOEVOQIIDKIKAFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPDFIAANOEERANNVIGTKOEVOQIIDKIKAFKATKVDKVVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDRNEAEISPTLYAIACWENVPIFGSPONTFVPGILDIAIANT 300
DB 241 VGLNDTMENLLAAVDRNEAEISPTLYAIACWENVPIFGSPONTFVPGILDIAIANT 300
QY 301 LIGGDDFKSGQTKMKSVLVDLVGAGIKPTISVYNHGNNDGMLSAPOFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDLVGAGIKPTISVYNHGNNDGMLSAPOFRSKEISKS 360
QY 361 NVDDMVNSNALIYEPGEHPDHVVYIKVPYVGSKRAMDEYSEIFMGKSTIYLAHTC 420
DB 361 NVDDMVNSNALIYEPGEHPDHVVYIKVPYVGSKRAMDEYSEIFMGKSTIYLAHTC 420
QY 421 EDSLAAPIIIDLVLLAEIETRIEFKAENEGKFSHPVATILSYLTAPLVPPTPVVN 480
DB 421 EDSLAAPIIIDLVLLAEIETRIEFKAENEGKFSHPVATILSYLTAPLVPPTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 12
US-10-025-003-12
;; Sequence 12, Application US/10025003
;; Publication No. US20030074685A1
;; GENERAL INFORMATION:
;; APPLICANT: Hitz, William
;; APPLICANT: Sebastian, Scott
;; APPLICANT: Grace, John
;; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
;; FILE REFERENCE: BB-1077-C
;; CURRENT APPLICATION NUMBER: US/10/025,003
;; PRIOR FILING DATE: 2002-05-07
;; PRIOR APPLICATION NUMBER: 08/835,751
;; PRIOR FILING DATE: APRIL 8, 1997
;; PRIOR APPLICATION NUMBER: PCT/US98/06822
;; PRIOR FILING DATE: APRIL 7, 1998
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 12
;; LENGTH: 510
;; TYPE: PRT
;; ORGANISM: Glycine max
US-10-025-003-12

Query Match 98.5%; Score 2593; DB 14; Length 510;

Best Local Similarity 98.4%; Pred. No. 2,1e-216;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60
DB 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60
QY 61 KLGWLVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGWLVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLPVMNPPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLPVMNPPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBERANVVKGTQOEVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240
DB 181 YDPDFIAANOEBERANVVKGTQOEVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240
QY 241 VGLNMTMENLAAVDRNEAISPSTLYAIACMENVPFINGSPONTFVPGGLIDLAIAART 300
DB 241 VGLNMTMENLAAVDRNEAISPSTLYAIACMENVPFINGSPONTFVPGGLIDLAIAART 300
QY 301 LIGDDPFSGQTKMSVLVDPLVAGAGIKPTSIYSYNHLGNNDGMMLSAPOTRSKESKS 360
DB 301 LIGDDPFSGQTKMSVLVDPLVAGAGIKPTSIYSYNHLGNNDGMMLSAPOTRSKESKS 360
QY 361 NVDDMVNSNALIYBEGHPDHVVVVKVPPYVDSKRAMDEYTSBIFMGKSTIVLHNTC 420
DB 361 NVDDMVNSNALIYBEGHPDHVVVVKVPPYVDSKRAMDEYTSBIFMGKSTIVLHNTC 420
QY 421 EDSLAAPIILDVLVLAELSTRIEPKAENEGKFSHPVATILSYLTKAPLVPGTTPVYN 480
DB 421 EDSLAAPIILDVLVLAELSTRIEPKAENEGKFSHPVATILSYLTKAPLVPGTTPVYN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 13

```

US-10-718-952-12
; Sequence 12, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-12

```

Query Match 98.5%; Score 2593; DB 16; Length 510;

Best Local Similarity 98.4%; Pred. No. 2,1e-216;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60
DB 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60

```

QY 61 KLGWLVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

DB 61 KLGWLVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

QY 121 IYAPFKSLPVMNPPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180

DB 121 IYAPFKSLPVMNPPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180

QY 181 YDPDFIAANOEBERANVVKGTQOEVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240

DB 181 YDPDFIAANOEBERANVVKGTQOEVOQIIKDIKAFKATKYDKVVLMTANTERYSNLV 240

QY 241 VGLNMTMENLAAVDRNEAISPSTLYAIACMENVPFINGSPONTFVPGGLIDLAIAART 300

DB 241 VGLNMTMENLAAVDRNEAISPSTLYAIACMENVPFINGSPONTFVPGGLIDLAIAART 300

QY 301 LIGDDPFSGQTKMSVLVDPLVAGAGIKPTSIYSYNHLGNNDGMMLSAPOTRSKESKS 360

DB 301 LIGDDPFSGQTKMSVLVDPLVAGAGIKPTSIYSYNHLGNNDGMMLSAPOTRSKESKS 360

QY 361 NVDDMVNSNALIYBEGHPDHVVVVKVPPYVDSKRAMDEYTSBIFMGKSTIVLHNTC 420

DB 361 NVDDMVNSNALIYBEGHPDHVVVVKVPPYVDSKRAMDEYTSBIFMGKSTIVLHNTC 420

QY 421 EDSLAAPIILDVLVLAELSTRIEPKAENEGKFSHPVATILSYLTKAPLVPGTTPVYN 480

DB 421 EDSLAAPIILDVLVLAELSTRIEPKAENEGKFSHPVATILSYLTKAPLVPGTTPVYN 480

QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

US-10-424-599-154864

```

; Sequence 154864, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154864
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1.pcp
US-10-424-599-154864

```

Query Match 95.9%; Score 2524; DB 15; Length 510;

Best Local Similarity 94.7%; Pred. No. 2,1e-210;
Matches 483; Conservative 18; Mismatches 9; Indels 0; Gaps 0;

```

QY 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60
DB 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60
QY 61 KLGWLVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGWLVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLPVMNPPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLPVMNPPDIFVGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180

```

QY 181 YDPDEIAANOEBRANNVTKGTKEQVOOIIKDIIKAFKATKVYKVYVMTANTERYSNLV 240
DB 181 YDPDEIAANOEBRANNVTKGTKEQVOOIIKDIIKAFKATKVYKVYVMTANTERYSNLV 240
QY 241 VGLANTMENTLAAVBRNEAEISPTLXAIACVMEVVPFINSPOPTFVPGIDLAIAANT 300
DB 241 VGLANTMENTLAAVBRNEAEISPTLXAIACVMEVVPFINSPOPTFVPGIDLAIAANT 300
QY 301 LIIGDDFKSGOTKMSVLVDPLVAGAIKPTISVSYNHLGNDGMNLSAPQTFRSKEISKS 360
DB 301 LIIGDDFKSGOTKMSVLVDPLVAGAIKPTISVSYNHLGNDGMNLSAPQTFRSKEISKS 360
QY 361 NVDDMVNSNALIYEPGEHPDHVVYIKVYPVGDSCRANDEYTSSEIFMGKSTIVLANTC 420
DB 361 NVDDMVNSNALIYEPGEHPDHVVYIKVYPVGDSCRANDEYTSSEIFMGKSTIVLANTC 420
QY 421 EDSLLAAPILDLVLLAEISTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPVNVN 480
DB 421 EDSLLAAPILDLVLLAEISTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPVNVN 480
QY 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510

RESULT 15
US-10-442-017-15

; Sequence 15, Application US/10442017
; Publication No. US20030215950A1
; GENERAL INFORMATION:
; APPLICANT: Battelle Memorial Institute
; APPLICANT: Lasure, Linda L.
; APPLICANT: Dai, Ziyu
; TITLE OF INVENTION: Isolated Polynucleotides and Methods of Promoting a Morphology In
; TITLE OF INVENTION: a Fungus
; FILE REFERENCE: BA4-195
; CURRENT APPLICATION NUMBER: US/10/442,017
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60/382,132
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 505
; TYPE: PRN
; ORGANISM: Sesamum indicum
US-10-442-017-15

Query Match 92.4%; Score 2432; DB 15; Length 505;

Best Local Similarity 92.3%; Pred. No. 2,1e-202;

Matches 466; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

QY 6 FVECEPNVYKTEETELOSVYVNETTELVEHNRNGTQWIVKPSVYKPEKTIHVPLGVM 65
DB 1 FVECEPNVYKTEETELOSVYVNETTELVEHNRNGTQWIVKPSVYKPEKTIHVPLGVM 65
QY 66 LVGMGNNSTLTGVIANREGISWATKDIQOANYFGSLTOASAIRVGSFOGBETIYAPF 125
DB 66 LVGMGNNSTLTGVIANREGISWATKDIQOANYFGSLTOASAIRVGSFOGBETIYAPF 125
QY 126 KSLLPVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKLRPYMESMLPLPGIYDPDF 185
DB 126 KSLLPVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKLRPYMESMLPLPGIYDPDF 185
QY 186 IAANOEERANNVTKGTKEQVOOIIKDIIKAFKATKVYKVYVMTANTERYSNLVGND 245
DB 186 IAANOEERANNVTKGTKEQVOOIIKDIIKAFKATKVYKVYVMTANTERYSNLVGND 245
QY 246 TWENTILAAVDRNEAEISPTLXAIACVMEVVPFINSPOPTFVPGIDLAIAANTLIGGD 305
DB 246 TWENTILAAVDRNEAEISPTLXAIACVMEVVPFINSPOPTFVPGIDLAIAANTLIGGD 305
QY 306 DFKSGOTKMSVLVDPLVAGAIKPTISVSYNHLGNDGMNLSAPQTFRSKEISKSNNVDD 365

DB 301 DFKSGOTKMSVLVDPLVAGAIKPTISVSYNHLGNDGMNLSAPQTFRSKEISKSNNVDD 360
QY 366 MNSNALIYEPGEHPDHVVYIKVYPVGDSCRANDEYTSSEIFMGKSTIVLANTCEDSL 425
DB 361 MNSNALIYEPGEHPDHVVYIKVYPVGDSCRANDEYTSSEIFMGKSTIVLANTCEDSL 420
QY 426 AAPILDLVLLAEISTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPVNVNALSQ 485
DB 421 AAPILDLVLLAEISTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPVNVNALSQ 480
QY 486 RAMLENTIRACVGLAPENNMLEYK 510
DB 481 RAMLENTIRACVGLAPENNMLEYK 505

Search completed: June 7, 2005, 17:03:07
Job time : 84.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:31:27 ; Search time 31 Seconds
(without alignments)
1228.098 Million cell updates/sec

Title: US-10-718-952-2
Perfect score: 2632
Sequence: 1 MEIENFKVCEPVKXKTETETI.....NIMBACVGLAPENNILEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2336	88.8	510	3	US-09-118-442-11 Sequence 11, Appl
2	2336	88.8	510	3	US-09-677-064-11 Sequence 11, Appl
3	2335	88.7	510	4	US-09-727-628-2 Sequence 2, Appl
4	1382.5	52.5	533	4	US-09-734-237B-73 Sequence 73, Appl
5	1382.5	52.3	534	4	US-09-734-237B-73 Sequence 73, Appl
6	1376.5	52.3	525	4	US-09-248-796A-17234 Sequence 17234, A
7	541	20.6	444	4	US-09-902-540-12518 Sequence 12518, A
8	113.5	4.3	730	4	US-09-107-433-4707 Sequence 4707, Ap
9	113.5	4.3	731	4	US-09-583-110-4720 Sequence 4720, Ap
10	112.5	4.3	842	4	US-09-489-039A-11570 Sequence 11570, A
11	111.5	4.2	452	4	US-09-248-796A-20645 Sequence 20645, A
12	111.5	4.2	731	1	US-09-014-897-2 Sequence 2, Appl
13	111.5	4.2	731	1	US-08-731-716-2 Sequence 2, Appl
14	111	4.2	935	4	US-09-512-250C-33 Sequence 33, Appl
15	109.5	4.2	747	4	US-09-583-110-3066 Sequence 3066, Ap
16	109.5	4.2	751	4	US-09-107-433-3375 Sequence 3375, Ap
17	108	4.1	2465	2	US-08-596-291-3 Sequence 3, Appl
18	108	4.1	2465	3	US-09-100-804-3 Sequence 3, Appl
19	108	4.1	2466	3	US-09-080-855-12 Sequence 12, Appl
20	108	4.1	2466	4	US-09-566-076-12 Sequence 12, Appl
21	108	4.1	2466	5	PCT-US94-09943-2 Sequence 2, Appl
22	108	4.1	2485	3	US-09-290-640-46 Sequence 46, Appl
23	108	4.1	2485	4	US-09-665-615B-46 Sequence 46, Appl
24	107.5	4.1	1363	4	US-09-583-110-5243 Sequence 5243, Ap
25	107.5	4.1	1367	4	US-09-107-433-4883 Sequence 4883, Ap
26	106.5	4.0	778	4	US-09-489-039A-8455 Sequence 8455, Ap
27	106	4.0	539	4	US-09-543-681A-6877 Sequence 6877, Ap

28	106	4.0	1242	4	US-09-107-532A-5241 Sequence 5241, Ap
29	103.5	3.9	912	4	US-09-248-796A-20032 Sequence 20032, A
30	101.5	3.9	2042	4	US-09-077-098A-6 Sequence 6, Appl
31	101	3.8	449	4	US-09-107-433-4132 Sequence 4132, Ap
32	101	3.8	454	4	US-09-583-110-3922 Sequence 3922, Ap
33	100.5	3.8	545	4	US-09-248-796A-14131 Sequence 14131, A
34	100.5	3.8	835	4	US-09-438-185A-728 Sequence 728, Ap
35	100.5	3.8	10382	3	US-09-134-001C-3159 Sequence 3159, Ap
36	99.5	3.8	390	3	US-08-745-995A-18 Sequence 18, Appl
37	99.5	3.8	390	4	US-09-005-352-18 Sequence 18, Appl
38	99.5	3.8	420	3	US-08-745-995A-14 Sequence 14, Appl
39	99.5	3.8	420	4	US-09-005-352-14 Sequence 14, Appl
40	99.5	3.8	1091	4	US-09-583-110-3823 Sequence 3823, Ap
41	99.5	3.8	1102	4	US-09-107-433-4859 Sequence 4859, Ap
42	99	3.8	389	4	US-09-489-039A-13228 Sequence 13228, A
43	99	3.8	414	4	US-09-107-433-3327 Sequence 3327, Ap
44	99	3.8	463	4	US-09-710-279-960 Sequence 960, App
45	99	3.8	551	4	US-09-489-039A-11049 Sequence 11049, A

ALIGNMENTS

```

RESULT 1
US-09-118-442-11
; Sequence 11, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-118-442-11
Query Match      88.8%; Score 2336; DB 3; Length 510;
Best Local Similarity 87.3%; Pred. No. 4.5e-223;
Matches 445; Conservative 31; Mismatches 34; Indels 0; Gaps 0;
1 MEIENFKVCEPVKXKTETETIIGSVNYETTELVENNRNGTYOIVKPKSVYEFKNIHVP 60
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1 MEISFRVSPHVRGPMIEBSERYDTTELVEHGDSRVRVBPVKVNFRTTAVP 60
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
61 KGVNVLVGVGNGGSLTLAGVIANEGISWATKDVQOQANYYGSLTHASTIVGSNGER 120
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
61 KGVNVLVGVGNGGSLTLAGVIANEGISWATKDVQOQANYYGSLTQASAIIVGSFOGEE 120
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
121 IYAPFKSLIPVNPDPDIVFGGWDINMMIADAMARAKFDIDLOKOLRPYMESMTPLPGI 180
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
121 IYAPFKSLIPVNPDPDIVFGGWDINMMIADAMARAKFDIDLOKOLRPYMESMTPLPGI 180
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
181 YDPDFIAANOEERANVVIKGTQEQVOQIINDIKAFKATKYDKVAVIWTANTERYSNLV 240
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
181 YDPDFIAANOEERANVVIKGTQEQVOQIINDIKAFKATKYDKVAVIWTANTERYSNLV 240
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
181 YDPDFIAANOEERANVVIKGTQEQVOQIINDIKAFKATKYDKVAVIWTANTERYSNLV 240
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
241 VGLNPTMENLAAVDRNBAEISPTLYAIACMENVPTINSQPTFPVGLIDLAIARNT 300

```



```
Db 241 AGANDTMENTLASVDKNEAEVSPSTLYAACWMEGVPIGNSPQNTFVPGILDLAIKNNC 300
Qy 301 LIGGDDFSGGQTKKMSVLDVFLVGAGIKPTSIIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFSGGQTKKMSVLDVFLVGAGIKPTSIIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALILYERGHDPHVVYIKVYPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVDDMVNSNALILYERGHDPHVVYIKVYPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAEISTRIEKAENEGKFSFHPVATILSYTKAPLVPPTPVVN 480
Db 421 EDSLAAPIIIDLVLLAEISTRIEKAENEGKFSFHPVATILSYTKAPLVPPTPVVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALAKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 2
US-09-677-064-11

```
/ Sequence 11, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
/ TITLE OF INVENTION: Plants and Uses Thereof
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677,064
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ PRIOR FILING DATE: 1998-07-17
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 11
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Zea mays
US-09-677-064-11
```

Query Match 88.8%; Score 2336; DB 3; Length 510;
Best Local Similarity 87.3%; Pred. No. 4.5e-223;
Matches 445; Conservative 31; Mismatches 34; Indels 0; Gaps 0;

```
Qy 1 MFIEHFKECPNVKKTETEISQVYNYETTELVHNRNGTYQWYVAKPKSVKYEFTKNIHVP 60
Db 1 MFIEHFKECPNVKKTETEISQVYNYETTELVHNRNGTYQWYVAKPKSVKYEFTKNIHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIYANREGISWATKDKIQOANYFGSLTQASATIRVSGFOGEE 120
Db 61 KLGWLVGMGNGNSTLTGVIYANREGISWATKDKIQOANYFGSLTQASATIRVSGFOGEE 120
Qy 121 IYAPFKSLPLPMVNPDDIVFGGWDISNNMLADAMARAKYFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFKSLPLPMVNPDDIVFGGWDISNNMLADAMARAKYFDIDLOKOLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANQSGRANNVIKGTKEQVEQIYKIDIRFKEKKNVYDKVWVMTANTERYSNLV 240
Db 181 YDPDFIAANQSGRANNVIKGTKEQVEQIYKIDIRFKEKKNVYDKVWVMTANTERYSNLV 240
Qy 241 VGLNDTMENTLASVDKNEAEVSPSTLYAACWMEGVPIGNSPQNTFVPGILDLAIKNNC 300
Db 241 VGLNDTMENTLASVDKNEAEVSPSTLYAACWMEGVPIGNSPQNTFVPGILDLAIKNNC 300
```

```
Qy 301 LIGGDDFSGGQTKKMSVLDVFLVGAGIKPTSIIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFSGGQTKKMSVLDVFLVGAGIKPTSIIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALILYERGHDPHVVYIKVYPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVDDMVNSNALILYERGHDPHVVYIKVYPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAEISTRIEKAENEGKFSFHPVATILSYTKAPLVPPTPVVN 480
Db 421 EDSLAAPIIIDLVLLAEISTRIEKAENEGKFSFHPVATILSYTKAPLVPPTPVVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALAKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 3
US-09-727-628-2

```
/ Sequence 2, Application US/09727628
/ Patent No. 6791013
/ GENERAL INFORMATION:
/ APPLICANT: Armstrong, Katherine
/ APPLICANT: Hey, Timothy D
/ APPLICANT: Folkerts, Otto
/ APPLICANT: Smith, Kelley A
/ APPLICANT: Hopkins, Nicole L
/ TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
/ FILE REFERENCE: 50597
/ CURRENT APPLICATION NUMBER: US/09/727,628
/ CURRENT FILING DATE: 2000-12-01
/ PRIOR APPLICATION NUMBER: US 60/168,612
/ PRIOR FILING DATE: 1999-12-02
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Zea mays
US-09-727-628-2
```

Query Match 88.7%; Score 2335; DB 4; Length 510;
Best Local Similarity 87.6%; Pred. No. 5.6e-223;
Matches 447; Conservative 29; Mismatches 34; Indels 0; Gaps 0;

```
Qy 1 MFIEHFKECPNVKKTETEISQVYNYETTELVHNRNGTYQWYVAKPKSVKYEFTKNIHVP 60
Db 1 MFIEHFKECPNVKKTETEISQVYNYETTELVHNRNGTYQWYVAKPKSVKYEFTKNIHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIYANREGISWATKDKIQOANYFGSLTQASATIRVSGFOGEE 120
Db 61 KLGWLVGMGNGNSTLTGVIYANREGISWATKDKIQOANYFGSLTQASATIRVSGFOGEE 120
Qy 121 IYAPFKSLPLPMVNPDDIVFGGWDISNNMLADAMARAKYFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFKSLPLPMVNPDDIVFGGWDISNNMLADAMARAKYFDIDLOKOLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANQSGRANNVIKGTKEQVEQIYKIDIRFKEKKNVYDKVWVMTANTERYSNLV 240
Db 181 YDPDFIAANQSGRANNVIKGTKEQVEQIYKIDIRFKEKKNVYDKVWVMTANTERYSNLV 240
Qy 241 VGLNDTMENTLASVDKNEAEVSPSTLYAACWMEGVPIGNSPQNTFVPGILDLAIKNNC 300
Db 241 VGLNDTMENTLASVDKNEAEVSPSTLYAACWMEGVPIGNSPQNTFVPGILDLAIKNNC 300
Qy 301 LIGGDDFSGGQTKKMSVLDVFLVGAGIKPTSIIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFSGGQTKKMSVLDVFLVGAGIKPTSIIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALILYERGHDPHVVYIKVYPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVDDMVNSNALILYERGHDPHVVYIKVYPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
```

APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

```

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17234
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17234

Query Match      52.3%; Score 1376.5; DB 4; Length 525;
Best Local Similarity 52.2%; Pred. No. 1e-127;
Matches 263; Conservative 98; Mismatches 128; Indels 15; Gaps 5;

QY 17 ETEISVYNYETTELIVHERNGTYOMIVKPKSVKYEFEKINIHVPKLGVMVGMGNGNST 76
DB 21 DDLTKTKTYENS-VVEDDANG--KFIVTPKSDYEFKVDLKVPRVGLLVIGGNGT 77
QY 77 LTGVIANREGISWATKXIQOANYFGSLTQASAIRVG--SFQSEIYAPFKSLIPVNP 134
DB 78 LLGATLADKHNISFENKEGVKPNYGSVTVQASTVKIGVDKETGBDVYVFPFNSIYPMVNP 137
QY 135 DDIVGCGDISNMNLADMAKAVPDIDLQKQRLPYMESMLPLPGIYDPDFIAAOERA 194
DB 138 NDLVVDGWDISGLPLDQAMKRAKVLDTVLQKQLYPLEKKPKLESIIYPDFIALQSEBA 197
QY 195 NNVIK-----GTRKQEOVOQIIKDIAKFKATKVDKVVLMNTANTERSYMLVVGNDT 246
DB 198 NNVEFQVNGEVKTDKMAKADVEKIRKDIRFKAKNELDKYIIIMTANTERYADLVNVDI 257
QY 247 MENLLAADRNEAESIPSTLYAIACMVENVPFINGSPOFTYVGLIDLAIAANTLIGDD 306
DB 258 ADNLKISIKESHESIELAPSTVFVAASILEKVPYINGSPQTFVPVGYELAEKYDSFICGD 317
QY 307 FRSQGTAKSVYVDLVNGIGIKPTISVYNHIGNNDGMKLSAPQFRSKETSKSNVNDM 366
DB 318 FRSQGTAKSVYAOFLVNDGIRKPLSIASYNHLGNNDGVNLSPKOFRSKETSKOSVVDI 377
QY 367 VNSNAILY--EPGEHPDHVVIVKVPYVDSKRAMDEYSEIFMGSKSTIVLHNTCESL 424
DB 378 ISSNELINKESGDVDCIVIKIIPAYDSKVADEYSSELMGHNKISIHNCESL 437
QY 425 LAAPILDLVLLAELSTRIEFAENEGKFSFHPVATIIISYTKAPLVPPGTPVNVNLSK 484
DB 438 LATPLILDLVATEFATRVQVKGPKSDYDELVPVASLSLWMLKAPLARPGFPIINGLAK 497
QY 485 ORAMLENTIRACVGLAPENNMILE 508
DB 498 ORQOLVNLISLVGLPIDNELRFE 521

RESULT 7
US-09-902-540-12518
; Sequence 12518, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12518
```

```

; LENGTH: 444
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12518

Query Match      20.6%; Score 541; DB 4; Length 444;
Best Local Similarity 29.8%; Pred. No. 8.3e-45;
Matches 140; Conservative 92; Mismatches 190; Indels 48; Gaps 12;

QY 52 EFKTNIHVP--KLGMVLVGMGNGNSTLTGVIANRREGISWATKDIQOANYFGSLTQAS 109
DB 2 ENKSVAKPDGLAVLPLGLAVS--TTLMAVELARKG-----KHPISGLTQMG 50
QY 110 AIRVGS--FQGEIYAPFKSLIPVNPDDIVEGCGDISNMNLADMAKAVPDIDLQKQLR 168
DB 51 TARGKRTDKRTV--KINELVPLAEIKVAFGADIIREDAYEVAVRGVLSDKLEEVK 108
QY 169 PYMESMLPLPGIYDPDFIAAOERANVIGTK--QGVQOQIIKDIAFKFATKVDKVV 226
DB 109 PPLQSIKPKQGVHDEPFRVIEA----NHKATKTHRESIALRQDIRDFKELNATRAV 164
QY 227 VLMNTANTERYSNLVVGLNDTMENTLLAAYDRNEAESIPSTLYAIACMVENVPFINGSPO-N 285
DB 165 MVVCSVETFPRLPESFK--TLAAFEKALDENSEPDINPTALTYAIAKEGVFPANAIPNAS 223
QY 286 TFPVGLIDLAIAARNTLIGDDPKSGQTKMSVLVDPLVAGAIKPTISVYNHLGNNDGMN 345
DB 224 VDTPLQGMARQESVAVAGRLKSGQTKMTVIALKARMLGLDGFSTNILGRDGEV 283
QY 346 LSAQTPFKSKISKSNVVDVNVNSNAILYEPGEHPD-----HYVVIKYVYVGSKRAM 399
DB 284 LDDPAFAFKAKVTYKSSVVDITL-----QDLYPELYKKAHKAHAIHYPPRGDAKGM 336
QY 400 DEYSEIMGKSTVLHNTCESLSLAAPITLDVLTLAELSTRIEFAKNEGKFSFHPV 459
DB 337 DNIDITGMVGPVQIKVFLCRDSTLAPLVLDIALFLDLAKRLEWREGIQEM----- 388
QY 460 ATILSYLTKAPLVPPGTPVNVNLSKORAMLENTIRACVGLAPENNMILEY 509
DB 369 ---MSFYKTPMAQGLPVEHDLFLQLTKMLKVLAVAGEPITHTLGDY 435

RESULT 8
US-09-107-433-4707
; Sequence 4707, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: Nucleic Acid and Amino Acid
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
```

NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4707:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURES:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...730
SEQUENCE DESCRIPTION: SEQ ID NO: 4707:
US-09-107-433-4707
Query Match 4.3%; Score 113.5; DB 4; Length 730;
Best Local Similarity 20.8%; Pred. No. 0.058;
Matches 108; Conservative 64; Mismatches 183; Indels 165; Gaps 29;
QY 90 WATKDKIQOANYFG-----SLTQASAIRVGSFGGEIYAPFKSLPMPVNPDIYFGMD 143
DB 218 WGEVDASKR--YFGVSASEVSLDQAATL--AGMLKGELYNPLNSVEDSTNRDVT----- 269
QY 144 ISNMNLA---DAMARAKVFDIDLQKQLRPMESMLP---LPGIYDPDFIAA-----NOE 191
DB 270 LQNMVAAGYIDKNGQTEAEVDMTSQLDHKYEGKISDYRYPSPYFPAVVAEAVSKVLTTEE 329
QY 192 EKAN--VIKTKQEQVQOIIKDIKAFKA-----TKYDKVYVLM 229
DB 330 EIVNNGYRIYTELDQNYQAMQIYVENISLFPRAEDGTPAQSVALBPKTGVRGVQ 389
QY 230 TANTER-----YSN-----LVV-----GNDTMENTLAADVREAEI 261
DB 390 VADNDKGTGRNRYATQSKSPGSTIKPLVYTPPAVEAGMALNKQDNTHTMQYDSYKV- 448
QY 262 SPSTLYALACVMEVNPFGSPONTFVPGLI---DLAIARNTLIGDDPKSGOTKRSYL 318
DB 449 ---NYAGIKTSREVMYQALAESLNLPAVATVNDLGVK-APEAGEKKGGLMKEKVDRL 503
QY 319 VDFVAGAGIK--PTSIIVSYNHLGNNDG-----NLAPQTFRSKEISKSNVVD 365
DB 504 -GVALGSGVETNPLQMAQVAAFAFANEGLMPEAFISRIENASG-QVIANHKNSQKRVIDK 561
QY 366 WY-----NSNALIYEPGEHPDHVVVYKVPYVGSKRAMD-EYTSSEIFMG- 410
DB 562 SVADKMTSMGLGTFNNGTGISSSPA---DYVMAK---TGTEAVFNEEYISDQWVICY 614
QY 411 -KSTIVLH-----NTCEDSLAAPILLDLVLLAELSTRIEFAENEGKFSHPVA-TTL 463
DB 615 TPDVVISHLGFPPTDENHYLAG-----STNGCAAHVRNRIANTIL 655
QY 464 SYLTKAPLVPGTPTVYNALSKORAMLENIMRACVGLAPEN 503
DB 656 -----PYTPGSTFTVENAYKON-----GIAPAN 678
RESULT 9
US-09-583-110-4720
Sequence 4720, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 4720
LENGTH: 731
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-4720
Query Match 4.3%; Score 113.5; DB 4; Length 731;
Best Local Similarity 20.8%; Pred. No. 0.058;
Matches 108; Conservative 64; Mismatches 183; Indels 165; Gaps 29;
QY 90 WATKDKIQOANYFG-----SLTQASAIRVGSFGGEIYAPFKSLPMPVNPDIYFGMD 143
DB 215 WGEVDASKR--YFGVSASEVSLDQAATL--AGMLKGELYNPLNSVEDSTNRDVT----- 266
QY 144 ISNMNLA---DAMARAKVFDIDLQKQLRPMESMLP---LPGIYDPDFIAA-----NOE 191
DB 267 LQNMVAAGYIDKNGQTEAEVDMTSQLDHKYEGKISDYRYPSPYFPAVVAEAVSKVLTTEE 326
QY 192 EKAN--VIKTKQEQVQOIIKDIKAFKA-----TKYDKVYVLM 229
DB 327 EIVNNGYRIYTELDQNYQAMQIYVENISLFPRAEDGTPAQSVALBPKTGVRGVQ 386
QY 230 TANTER-----YSN-----LVV-----GNDTMENTLAADVREAEI 261
DB 387 VADNDKGTGRNRYATQSKSPGSTIKPLVYTPPAVEAGMALNKQDNTHTMQYDSYKV- 445
QY 262 SPSTLYALACVMEVNPFGSPONTFVPGLI---DLAIARNTLIGDDPKSGOTKRSYL 318
DB 446 ---NYAGIKTSREVMYQALAESLNLPAVATVNDLGVK-APEAGEKKGGLMKEKVDRL 500
QY 319 VDFVAGAGIK--PTSIIVSYNHLGNNDG-----NLAPQTFRSKEISKSNVVD 365
DB 501 -GVALGSGVETNPLQMAQVAAFAFANEGLMPEAFISRIENASG-QVIANHKNSQKRVIDK 558
QY 366 WY-----NSNALIYEPGEHPDHVVVYKVPYVGSKRAMD-EYTSSEIFMG- 410
DB 559 SVADKMTSMGLGTFNNGTGISSSPA---DYVMAK---TGTEAVFNEEYISDQWVICY 611
QY 411 -KSTIVLH-----NTCEDSLAAPILLDLVLLAELSTRIEFAENEGKFSHPVA-TTL 463
DB 612 TPDVVISHLGFPPTDENHYLAG-----STNGCAAHVRNRIANTIL 652
QY 464 SYLTKAPLVPGTPTVYNALSKORAMLENIMRACVGLAPEN 503
DB 653 -----PYTPGSTFTVENAYKON-----GIAPAN 675
RESULT 10
US-09-489-039A-11570
Sequence 11570, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Bretton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 11570
LENGTH: 842
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11570

Query Match 4.3%; Score 112.5; DB 4; Length 842;
Best Local Similarity 22.9%; Pred. No. 0.092;
Matches 100; Conservative 63; Mismatches 139; Indels 135; Gaps 27;

QY 30 ELVHENRNGTWMYIKPKSVKTEFKNTNHPKLGMYLWGWGNNSTLTCGVIANREGIS 89
DB 426 QLVDDOVQDATAVYL--KGICPD-----NPFQGDFFINAFYREGT----- 463
QY 90 WATKXKIOQANVFGLTQASAIRVGSFQGEELIYAFPKSLPMVNPDDIVFGWMDISNM-- 147
DB 464 -----QRFNDFENLRVAGQY-IGSF--TMTPTASTQQYVS-----DLKNOVG 502
QY 148 NLADAMAPAKYVD-----IDLQKOLRPYMESMLPLPGIY-----DPDEFIANOEER 193
DB 503 NGEGLAGAAAFADHYQAAQRILISQKSDPF--QSAVDI-GAYKPISTNPDAIABEVNR 560
QY 194 --ANNVIGK-----TKQEVQOIIKDIFKFKATKDKVVLMTANTERYSNLVVGL 243
DB 561 YAAQDOLKAIGITPPLSKOE--SQVLTLD--AVRSTVDVNCAL-----SLQGL 605
QY 244 NDTM-ENLLAAYDRNEAISPSTLYALCV-MENVPIINGS--PONTFV--PGIIDL 295
DB 606 GTLPKALRSYASSIAPGSPGTAVALLLGQDDQYDRSGIITPSQFVSXKPTLDDKD 665
QY 236 IARNTLIGDDPKSGQTKKSVLVDPLVAGIKPTSIYSYNHL-----GNNDGMNLS 347
DB 666 VAK-TVLAGDQMLNPTKMKD-----AGISAVSIPDEKLTQYFDKEVGNAPAYSQ 716
QY 348 APQ-----TFPSKEISKSNNV--DDMNNSNALIYEPGE-----HPHYVVIK 387
DB 717 ARQVAMGNFRSAYALAYQSGDASKTNTVSPSDIAEKAAQYATGAVYKGLGSDV-- 773
QY 388 YVPYVGDGSKRAMDEXTS 404
DB 774 -MPFGMDKTTFRDRYTA 789

RESULT 11
US-09-248-796A-20645
; Sequence 20645, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20645
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20645

Query Match 4.2%; Score 111.5; DB 4; Length 452;
Best Local Similarity 20.6%; Pred. No. 0.041;
Matches 105; Conservative 83; Mismatches 157; Indels 165; Gaps 31;

QY 11 EVVXKTER---EIGSYVYNE-TTELVH---ENRNGYQWIVPKSVKYEFK----- 54
DB 2 PSIVVPDSTTKQYQATNNSENATEEYPIIESEK-----IVLMDAFNYLTKIILLOSIVAN 55
QY 55 -TNIHVPKLGWLV-----GMGNGNGSTLTCGVIANREGISATKDKIOQANV-FGSLTQA 108
DB 56 HPTITINQIMLLIIPSLTWSROSIEYITKYIENLEITAFNIIIDLSLAATGVCQLTNS 115
QY 109 SAIRVGSFQGEELIYAFPKSLPMVNPDDIVFGWMDISN-----MNL-----A 150

DB 116 TVVVY---DDENI-----QIVPVVGYQAIKPKAGKLINKEGSIITTSREIKONLPIUTSQOI 167
QY 151 DAMARAKYVD--IDLQKOLRPYMESMLPLPGIYPPDF-----IANOEERANNVYKGTQ 203
DB 168 EDLKNSDIFEEVVIDQGVNLDYIKDITKTNNDEDENEFPAIVENQNGIEAIIISNPT 227
QY 204 EQVQOIIKDI-KAFKATK--VDKVV--VLMTANTERYSNLVVGLNDJTMENLLAAVDNR 257
DB 228 EQQKEQDQSNKPKNELEKKNFYIDSKTEKIIWIGK-ERFS-----GTN-----N 270
QY 258 EAELSPSTLYAIACMENVPIINGSPOPTFVPGIIDLAIANTLIGDDPKSGQTKMSV 317
DB 271 LVKLSSSIYS--SLSTIPID-KRQDCY-----DNIIIVGSIFFTPGLK-EAI 315
QY 318 LV---DPLV-----GAGI-----KPTSIYSYNHLGNNDGMNLSAPQTERS 354
DB 316 LIKANQDYLVAEPNADIOSINDPGVNTALIKYQOSTTINDINEGSGDNNN----- 366
QY 355 KEISKSNVDDMVNSNALIYEPGEHPDHVVVIXYVPYVGDGSKRAMDE----- 401
DB 367 ---SNSN-----SNSNQV-----PNSIKLVKYPDYPEPKKPEKEKGSWHDYFLGQ 411
QY 402 -YTSIEFMG-----GKSTIVLHNTCEDSL 425
DB 412 IYSKQIYSGSHHGKELFVSGDMYERVL 441

RESULT 12
US-09-014-897-2
; Sequence 2, Application US/09014897
; Patent No. H002085
; GENERAL INFORMATION:
; APPLICANT: Hoskins, JoAnn
; Jaskunas, S. Richard
; Rokey, Pamela K.
; Zhao, Genshi
; Rostek, Paul R. Jr.
; No. H002085rfs, Franklin H.
; TITLE OF INVENTION: Penicillin Binding Protein From
; Streptococcus Pneumoniae
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/014,897
; FILING DATE: 28-Jan-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,716
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-10,887
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

LENGTH: 935
TYPE: PRT
ORGANISM: mus. p.
US-09-512-250C-33

Query Match 4.2%; Score 111; DB 4; Length 935;

Best Local Similarity 18.9%; Pred. No. 0.15;

Matches 100; Conservative 76; Mismatches 218; Indels 134; Gaps 19;

```
QY 52 EKKTHIHFKLVGMVGMGN-----GSTLTG-----VIANREGISMARDK 94
DB 177 EPEPDQFSNLEALLIQIGPKCVLPGETTGMKGLROVIORGILLTERRADPSTKD 236
QY 95 KIQQANYFGSLTQASAIRVGSFQGEIYAPFKSLPMVN-----PDDIVFGMDISN--- 146
DB 237 IYQDLNRLKKGKQGINSAALPEMENQAVSSLSANVTFLELLSDDSFGFELATPDE 296
QY 147 ---NNIADAMARA-----KVFIDLQKQLRPYMESMLPLPG-----IYDPDFIAANQ 190
DB 297 SQYMLDMAAVARALNLFQGSVEDTGSGLAALLNCKCTAQCGRLVNQWIKQPLMDRRRI 356
QY 191 EERANNVIGKTQEOVOQIIRK--IKAPKEATKVDKVVVLTANTERYSNLVVGNDTME 248
DB 357 EERLNLVFAFVEDSELROSLQEDLLRRFPDLNRLAKKQFORQANLQDCYRLQGINQ--LP 415
QY 249 NLLAADRNEAEISPTLYAIACV-----MENVPFINGSPO 284
DB 416 SVIQALEKTEGR--HQALLAVFVTPILDRSPSKQEMIEFTLMDQVENHEFL---VK 471
QY 285 NTFVPGILDLAIARNTLIGDDPKSGQTKMKSVLVDPLVAGIKPTSTIVSYNHLGNNDGM 344
DB 472 PEFDPNLSLREVMGDL-----EKKQSTLINAARGLIDP-----GKQIKL 513
QY 345 NLSAP-----QTFRSKEISKSN-----VDDMVNSNALIYERGEHPDHVVVVIKVPYVD 394
DB 514 DSSAGFYFRVYTCCKEKKLRNNKNFSTVD--IQKGVFTNSELS--LNEEYTKNKE 569
QY 395 SKRADEYTSEIFMGKSTIVLHNTCEDSLAPILDLVLAELESTRIEFKANEKGKH 454
DB 570 YEAQDALYKEIVNISSGVPEMQTLND-----VLAHLDALYSF----- 608
QY 455 SPHPVATILSYLTAKPLVPFGTPVVNALSKORAMLENTIRACVGLAPE 502
DB 609 -----AHVSNAPVYVRPVLLEKGRILLKASHACVEVODE 647
```

RESULT 15

US-09-583-110-3066

/ Sequence 3066, Application US/09583110

/ Patent No. 6699703

/ GENERAL INFORMATION:

/ APPLICANT: Lynn Doucette-Stamm et al.

/ TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

/ TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics

/ FILE REFERENCE: PATH00-07A

/ CURRENT APPLICATION NUMBER: US/09/583,110

/ PRIORITY FILING DATE: 2000-05-26

/ PRIORITY FILING DATE: US 09/107,433

/ PRIORITY FILING DATE: 1998-06-30

/ PRIORITY FILING DATE: US 60/085,131

/ PRIORITY FILING DATE: 1998-05-12

/ PRIORITY FILING DATE: US 60/051,553

/ PRIORITY FILING DATE: 1997-07-02

/ NUMBER OF SEQ ID NOS: 5322

/ SEQ ID NO 3066

/ LENGTH: 747

/ TYPE: PRT

/ ORGANISM: Streptococcus pneumoniae

US-09-583-110-3066

Query Match 4.2%; Score 109.5; DB 4; Length 747;
Best Local Similarity 21.3%; Pred. No. 0.15;
Matches 102; Conservative 56; Mismatches 161; Indels 159; Gaps 25;

```
QY 2 FIENFKVECPNVKYTEIEIOGVNYETTELVENRNGTQMIIVKRSKYEFKTHIHVK 61
DB 173 FIENFGVEHTTVSDIQNSIVTFVVDHIDIFRLN-----KLNIDISE 214
QY 62 LGVMVGMGNNGSTLTGCVIANREGISMATKDKIQQANYFGSLTQASAIRVGSFQGEI 121
DB 215 ANIM-----NQTT---VLLDEKNIEL---LSKAPYLVS-----I 244
QY 122 YAPFKSLPMVNPDDIVFGMDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGIY 181
DB 245 VEDFSKL---SVDDFSLDNDL--KINLPSMANEPPVGVITLFDKRYFENEMVEYHDFV 299
QY 182 DPDFIAANO-----ERAN---NYIKGTQEOVOQI-----IKDIAF 216
DB 300 SPDISKSDQYKHGTAVTSLVLDGANLNPDLDDCGGNFRVHFGVSLQSGFNSFTIIKQI 359
QY 217 KEATKVDKVVVLTANTERYSNLVVGNDTW--ENLLAAVDR--NEAEISPTLYAIACVM 273
DB 360 KEIVSQNADIKVM-----NLSIGSNDEIRENFISAEGALLDEIOPENDVIFIIAGT- 410
QY 274 ENVPFINGSPONTFVPGILDLAIARNTLI--GGDF--KSGQTKMKSVLVDPLVAGIKPT 330
DB 411 -NASVINKRRKIGAP---ADSLNSTIIVNSVDPNNQSVSYSRGIVLSFFV---KPD 460
QY 331 SIVSYNHLGNND-----GNLSPAPQTFR-----SKEISKSNVVD- 364
DB 461 --VSYYGGNDDFINVCPPLGIRVAGTSFAAPFIARGAVLIHMGLSREBAKALLIDA 518
QY 365 -----DM-----VNSNALIYERGEHPDHVVVVIKVPYVGDSSKRAMEY 402
DB 519 AIPWMDKKTFTDLSLIGNGIVPIKMDILSTPDDE-----IKFI--VSDISRAYDTY 568
```

Search completed: June 7, 2005, 16:48:53
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 03:03:45 ; Search time 1094.05 Seconds

(without alignments)
9972.402 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760

Sequence: 1 cctctcttctctcttcttgc.....aatgtagctaatcttgcgt 1760

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10I_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11A_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1760	100.0	1760	US-10-025-003-1	Sequence 1, Appli
2	1760	100.0	1760	US-10-718-952-1	Sequence 1, Appli
3	1745.8	99.2	1989	US-10-424-599-70167	Sequence 70167, A
4	1533	87.1	1533	US-10-025-003-9	Sequence 9, Appli
5	1533	87.1	1533	US-10-718-952-9	Sequence 9, Appli
6	1531.4	87.0	1533	US-10-025-003-5	Sequence 5, Appli
7	1531.4	87.0	1533	US-10-718-952-5	Sequence 5, Appli

8	1465.8	83.3	1533	14	US-10-025-003-13	Sequence 13, Appl
9	1465.8	83.3	1533	14	US-10-025-003-15	Sequence 15, Appl
10	1465.8	83.3	1533	19	US-10-718-952-13	Sequence 13, Appl
11	1465.8	83.3	1533	19	US-10-718-952-15	Sequence 15, Appl
12	1464.2	83.2	1533	14	US-10-025-003-11	Sequence 11, Appl
13	1464.2	83.2	1533	14	US-10-718-952-11	Sequence 11, Appl
14	1296.4	73.7	2018	18	US-10-424-599-12022	Sequence 12022, A
15	1136.8	64.6	2582	18	US-10-424-599-12021	Sequence 12021, A
16	1007	57.2	1324	19	US-10-767-701-14714	Sequence 14714, A
17	1000.6	56.9	1941	20	US-10-739-930-2091	Sequence 2091, Ap
18	999.2	56.8	2576	19	US-10-437-963-70769	Sequence 70769, Ap
19	995.8	56.6	2284	20	US-10-425-115-47189	Sequence 47189, A
20	995.4	56.6	1533	9	US-09-938-842A-1438	Sequence 1438, Ap
21	995.4	56.6	1533	11	US-09-938-842A-1438	Sequence 1438, Ap
22	994.6	56.5	2582	18	US-10-424-599-12021	Sequence 12021, A
23	994.2	56.5	1879	18	US-10-425-114-5704	Sequence 5704, Ap
24	994.2	56.5	1912	18	US-10-425-114-5483	Sequence 5483, Ap
25	994.2	56.5	1914	18	US-10-425-114-5612	Sequence 5612, Ap
26	994.2	56.5	1959	9	US-09-727-628-1	Sequence 1, Appli
27	994.2	56.5	1997	20	US-10-425-115-47190	Sequence 47190, A
28	994.2	56.5	2156	18	US-10-425-114-25858	Sequence 25858, A
29	993.8	56.5	2309	20	US-10-425-115-47194	Sequence 47194, A
30	992.6	56.4	1971	18	US-10-425-114-26983	Sequence 26983, A
31	992.6	56.4	2233	20	US-10-425-115-47187	Sequence 47187, A
32	991	56.3	2444	20	US-10-425-115-122151	Sequence 122151, A
33	990.4	56.3	1898	18	US-10-425-114-7081	Sequence 7081, Ap
34	984	55.9	1931	9	US-09-921-232-10	Sequence 10, Appl
35	984	55.9	1931	9	US-09-921-330-10	Sequence 10, Appl
36	984	55.9	1931	9	US-09-921-329-10	Sequence 205, App
37	956.4	54.3	1759	19	US-10-363-8829-205	Sequence 47180, A
38	920.8	52.3	1673	20	US-10-425-115-47180	Sequence 14009, A
39	919	52.2	1858	18	US-10-425-114-14009	Sequence 5864, Ap
40	919	52.2	2050	18	US-10-425-114-1864	Sequence 18524, A
41	917.4	52.1	1853	18	US-10-425-114-16524	Sequence 83553, A
42	917.4	52.1	2130	20	US-10-425-115-83553	Sequence 47182, A
43	906.4	51.5	1789	20	US-10-425-115-47182	Sequence 85785, A
44	904.6	51.4	2196	19	US-10-437-963-85285	Sequence 5619, Ap
45	676.6	38.4	1240	18	US-10-425-114-5619	

ALIGNMENTS

RESULT 1
US-10-025-003-1
; Sequence 1, Application US/10025003
; Publication No. US2003007465A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 9, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-1
Query Match 100.0%; Score 1760; DB 14; Length 1760;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTTCTTTATTCCTTTGTAATTCATTCATTCCTTTAACTTTGTGAAAAATATGTCA 60
1 CTTCTTTATTCCTTTGTAATTCATTCATTCCTTTAACTTTGTGAAAAATATGTCA 60
61 TCGAATATTTTAAAGTTGAGTGTCTTAATGTAGTACACGAGCTGAGATTCTGCG 120
61 TCGAATATTTTAAAGTTGAGTGTCTTAATGTAGTACACGAGCTGAGATTCTGCG 120
121 TGTACAACTACGAAACCAACGAACTTTGTCAGAGAACGAAATGGACCTACAGTGA 180
121 TGTACAACTACGAAACCAACGAACTTTGTCAGAGAACGAAATGGACCTACAGTGA 180
181 TTGTCAAAACCAAAATCTGTCAAAATAGAAATTTTAAACCAATCCATGTCTCTAAATAG 240
181 TTGTCAAAACCAAAATCTGTCAAAATAGAAATTTTAAACCAATCCATGTCTCTAAATAG 240
241 GGGTAAATGCTTGTGGGTTGGGGTGGAAACAAAGGCTCAACCTTCAACGGTGTATTG 300
241 GGGTAAATGCTTGTGGGTTGGGGTGGAAACAAAGGCTCAACCTTCAACGGTGTATTG 300
301 CTAAACGAGAGGGAATTTTCATGGGCTACAAAGACAAAGATTCAACAAAGCAATTCCTTG 360
301 CTAAACGAGAGGGAATTTTCATGGGCTACAAAGACAAAGATTCAACAAAGCAATTCCTTG 360
361 GCTCCCTCAACCAAGCCTTCAAGCTATCCAGATTGGGCTCTTCCAGGAGAGAAATCTATG 420
361 GCTCCCTCAACCAAGCCTTCAAGCTATCCAGATTGGGCTCTTCCAGGAGAGAAATCTATG 420
421 CCCCATTCAAGGCTGCTTCCAAATGTTAAACCTTCAAGCACTTGTGTTGGGGATGG 480
421 CCCCATTCAAGGCTGCTTCCAAATGTTAAACCTTCAAGCACTTGTGTTGGGGATGG 480
421 CCCCATTCAAGGCTGCTTCCAAATGTTAAACCTTCAAGCACTTGTGTTGGGGATGG 480
481 AATATGACAACTGAACTGAGTATCCAGTATCCAGGAGCAAGGTTTGAATCGATT 540
481 AATATGACAACTGAACTGAGTATCCAGTATCCAGGAGCAAGGTTTGAATCGATT 540
481 AATATGACAACTGAACTGAGTATCCAGTATCCAGGAGCAAGGTTTGAATCGATT 540
541 TGCAGAAAGAGTTGAGGCTTATCATGGAATCCATGCTTCCATCCCGGAATCTATGACC 600
541 TGCAGAAAGAGTTGAGGCTTATCATGGAATCCATGCTTCCATCCCGGAATCTATGACC 600
601 CGGATTTTCACTGCTGCAACCAAGAGAGCTGCTCAACAGCTCATCAAGGCAACAAAGC 660
601 CGGATTTTCACTGCTGCAACCAAGAGAGCTGCTCAACAGCTCATCAAGGCAACAAAGC 660
601 CGGATTTTCACTGCTGCAACCAAGAGAGCTGCTCAACAGCTCATCAAGGCAACAAAGC 660
661 AAGACAAAGTTCAACAAATCATCAAAAGACATCAAGGCGTTTAAAGGACCAAAAGTGG 720
661 AAGACAAAGTTCAACAAATCATCAAAAGACATCAAGGCGTTTAAAGGACCAAAAGTGG 720
721 ACAAGGTGTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 780
721 ACAAGGTGTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 780
781 TTAATGACCACTTGAAGATCTTGTGCTGTGTGGAAGAAATGAGGTGATTTCTC 840
781 TTAATGACCACTTGAAGATCTTGTGCTGTGTGGAAGAAATGAGGTGATTTCTC 840
841 CTTCCACCTTGAATGCAATGCTTGTGTTAAGAAATGTTCTTTCAATTAATGAAGCC 900
841 CTTCCACCTTGAATGCAATGCTTGTGTTAAGAAATGTTCTTTCAATTAATGAAGCC 900
901 CTACAGAACTTTTGTATGCAAGGCTGATTTGATCTTGCCATCCGAGGAACCTTTGATTTG 960
901 CTACAGAACTTTTGTATGCAAGGCTGATTTGATCTTGCCATCCGAGGAACCTTTGATTTG 960
961 GTGAGATGATCTTCAAGAGTGTGCAAGCAAAATGAAATCTGTGTTGTTGTTCTTGG 1020
961 GTGAGATGATCTTCAAGAGTGTGCAAGCAAAATGAAATCTGTGTTGTTGTTCTTGG 1020
1021 TGGGGGCTGGTATCAAGCAATCATATGATGATTAACCAATCTGGGAAACCAATGATG 1080
1021 TGGGGGCTGGTATCAAGCAATCATATGATGATTAACCAATCTGGGAAACCAATGATG 1080
1081 GTATGATCTTTCGGCTCCAAACTTTCGTTCCAGGAAATCTTCAAGAGCAAGCTTG 1140

1081 GTATGATCTTTCGGCTCCAAACTTTCGTTCCAGGAAATCTTCAAGAGCAAGCTTG 1140
1141 TTGATGATGATGTAACGAAATGCAATCCCTTGTAGGCTGGTGAACATCCAGACATG 1200
1141 TTGATGATGATGTAACGAAATGCAATCCCTTGTAGGCTGGTGAACATCCAGACATG 1200
1201 TTGTTGTTAATTAAGTATGCTTACGTAGGAGGAGCAGCAAGAGCCATGATGATGACA 1260
1201 TTGTTGTTAATTAAGTATGCTTACGTAGGAGGAGCAGCAAGAGCCATGATGATGACA 1260
1261 CTTACAGATATTTCAATGGTGGGAAAGAGACATTTGTTTTCACAAACATGCGAGATT 1320
1261 CTTACAGATATTTCAATGGTGGGAAAGAGACATTTGTTTTCACAAACATGCGAGATT 1320
1321 CCGCTTGTAGCTGCTCTTATCTTGTGACCTTGTCTTGTGAGCTCAGACTAGAA 1380
1321 CCGCTTGTAGCTGCTCTTATCTTGTGACCTTGTCTTGTGAGCTCAGACTAGAA 1380
1381 TCGAGTTTAAAGCTGAAATGAGGAAATTTCCACTATTCACCCAGTTGCTACATCC 1440
1381 TCGAGTTTAAAGCTGAAATGAGGAAATTTCCACTATTCACCCAGTTGCTACATCC 1440
1441 TCACTTACCTTCAACCAAGCTCTCTGCTTCCACCGGGTACACAGTGTGATGATGAT 1500
1441 TCACTTACCTTCAACCAAGCTCTCTGCTTCCACCGGGTACACAGTGTGATGATGAT 1500
1501 CAAAGCAGCTGCAATGCTGGAACCAATATGAGGAGCTTGTGATGATGAGCCAGAGA 1560
1501 CAAAGCAGCTGCAATGCTGGAACCAATATGAGGAGCTTGTGATGATGAGCCAGAGA 1560
1561 AATACATGATCTCGAGTACCAAGTGAAGCATGAGGACCGAAGAAATATATGTTGGGATG 1620
1561 AATACATGATCTCGAGTACCAAGTGAAGCATGAGGAGCGAAGAAATATATGTTGGGATG 1620
1621 CCTAGCTGAATGTTTATGTTAATATATATGTTGCTTAAATTTTGGCAAGGTATGAA 1680
1621 CCTAGCTGAATGTTTATGTTAATATATATGTTGCTTAAATTTTGGCAAGGTATGAA 1680
1681 TGCATCACTTCAATTAATGCTTGAAGGAGGAGCATATCTGTTTACTAGGAACATGAATG 1740
1681 TGCATCACTTCAATTAATGCTTGAAGGAGGAGCATATCTGTTTACTAGGAACATGAATG 1740
1741 AATGATGATATATTTGTGT 1760
1741 AATGATGATATATTTGTGT 1760

RESULT 2
US-10-718-952-1
; Sequence 1, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-1

Query Match 100.0%; Score 1760; DB 19; Length 1760;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTTTATTCCTTTTGTATTTTCAATTCATTCTTAATCTTTGTGAAAAATATATGTC 60
DB 1 CTCCTTTTATTCCTTTTGTATTTTCAATTCATTCTTAATCTTTGTGAAAAATATATGTC 60
QY 61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATTCACTCCG 120
DB 61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATTCACTCCG 120
QY 121 TGTACAACTACGAAACCAACCGAACTTTGTCAGAGAAACGGAATGGCACTTACAGTGA 180
DB 121 TGTACAACTACGAAACCAACCGAACTTTGTCAGAGAAACGGAATGGCACTTACAGTGA 180
QY 181 TTGTCAAAACCAAAATCTGTCAAAATAGCAATTTAAAAACAATCCATGTTCTTAAATTAG 240
DB 181 TTGTCAAAACCAAAATCTGTCAAAATAGCAATTTAAAAACAATCCATGTTCTTAAATTAG 240
QY 241 GGGTAAATGCTTGTGGGTGGGTGGAAACAAGGCTCAACCTCACCGGTGTGTTATTG 300
DB 241 GGGTAAATGCTTGTGGGTGGGTGGAAACAAGGCTCAACCTCACCGGTGTGTTATTG 300
QY 301 CTAAACGAGAGGCAATTTCAATGGGCTACAAAGGACAAATTTCAAGCCAAATTTACTTTG 360
DB 301 CTAAACGAGAGGCAATTTCAATGGGCTACAAAGGACAAATTTCAAGCCAAATTTACTTTG 360
QY 361 GCTCCCTCAACCAAGGCTCAGATCTCGAGTTGGGTCTCTTCAAGGAGAGAAATCTATG 420
DB 361 GCTCCCTCAACCAAGGCTCAGATCTCGAGTTGGGTCTCTTCAAGGAGAGAAATCTATG 420
QY 421 CCCCAATTCAGAGACCTGTCTTCAATGTTAACTTGAACGATTTGGGGAGTGGG 480
DB 421 CCCCAATTCAGAGACCTGTCTTCAATGTTAACTTGAACGATTTGGGGAGTGGG 480
QY 481 ATATCGCAACATGAACCTGGCTGATGCCATGGCCAGGCGAAAGGTTTGAATGATG 540
DB 481 ATATCGCAACATGAACCTGGCTGATGCCATGGCCAGGCGAAAGGTTTGAATGATG 540
QY 541 TCGAGAGAGGTTGAGGCTTATCAATGGAATCCATGCTTCACTCCCGGAATCTATGACC 600
DB 541 TCGAGAGAGGTTGAGGCTTATCAATGGAATCCATGCTTCACTCCCGGAATCTATGACC 600
QY 601 CGGATTTCAATGTGTCCAACCAAGAGAGGCTGCCAAACGTCATCAAGGGGCAAAAG 660
DB 601 CGGATTTCAATGTGTCCAACCAAGAGAGGCTGCCAAACGTCATCAAGGGGCAAAAG 660
QY 661 AAGAGCAAGTTCAACAAATCATTAAGACATCAAGGCTTTAAGAAAGCCAAAGTGG 720
DB 661 AAGAGCAAGTTCAACAAATCATTAAGACATCAAGGCTTTAAGAAAGCCAAAGTGG 720
QY 721 ACAAGTGTGTACTGTGGAAGTCCAAACAGAGAGTACAGATTTTGTGTGGGGCC 780
DB 721 ACAAGTGTGTACTGTGGAAGTCCAAACAGAGAGTACAGATTTTGTGTGGGGCC 780
QY 781 TTAATGACACATGAGAAATCTTCTGCTGTGTGACAGAAATGAGGCTGAGATTTCTC 840
DB 781 TTAATGACACATGAGAAATCTTCTGCTGTGTGACAGAAATGAGGCTGAGATTTCTC 840
QY 841 CTTCACCTTTGATGTCATTTGCTTGTGTTATGAAAAATGTTCTTTCATTAATGAAAGCC 900
DB 841 CTTCACCTTTGATGTCATTTGCTTGTGTTATGAAAAATGTTCTTTCATTAATGAAAGCC 900
QY 901 CTCAGAAACCTTTTGTACAGAGGCTGATGATCTTGCAATGCGAGAGAACCTTTGATTG 960
DB 901 CTCAGAAACCTTTTGTACAGAGGCTGATGATCTTGCAATGCGAGAGAACCTTTGATTG 960
QY 961 GTGAGATGATCTTCAAGAGTGTGTCAGACCAAAATGAAATCTGTGTGTTGATTTCTTG 1020
DB 961 GTGAGATGATCTTCAAGAGTGTGTCAGACCAAAATGAAATCTGTGTGTTGATTTCTTG 1020

QY 1021 TGGGGGCTGTATTCAGACCAATCATATAGTACGTTAACACCATCTGGGAAACAATGATG 1080
DB 1021 TGGGGGCTGTATTCAGACCAATCATATAGTACGTTAACACCATCTGGGAAACAATGATG 1080
QY 1081 GATGATCTTTCCGGCTCCACAACTTTCCGTTCCAGAGAAATCTTCCAGAGCAAGTTG 1140
DB 1081 GATGATCTTTCCGGCTCCACAACTTTCCGTTCCAGAGAAATCTTCCAGAGCAAGTTG 1140
QY 1141 TTGATGATATGTGTCAACAGCAATGCCATCTCTATGAGGCTGTGAACATCCAGACATG 1200
DB 1141 TTGATGATATGTGTCAACAGCAATGCCATCTCTATGAGGCTGTGAACATCCAGACATG 1200
QY 1201 TTGTTGTTAATTAAGTATGAGTCCCTTACGTAGGGGACAGCAAGAGGACATGATGATCA 1260
DB 1201 TTGTTGTTAATTAAGTATGAGTCCCTTACGTAGGGGACAGCAAGAGGACATGATGATCA 1260
QY 1261 CTTCAGAGATATTCATGGGTGAAAGAGCACCATTTGTCACAAACATCGAGAGATT 1320
DB 1261 CTTCAGAGATATTCATGGGTGAAAGAGCACCATTTGTCACAAACATCGAGAGATT 1320
QY 1321 CCTCTTACCTGCTCTATTAATCTTGAATGCTGCTTTTGTGAGCTCAGACTAGAA 1380
DB 1321 CCTCTTACCTGCTCTATTAATCTTGAATGCTGCTTTTGTGAGCTCAGACTAGAA 1380
QY 1381 TCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATTCACACCAAGTTGCTACCATCC 1440
DB 1381 TCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATTCACACCAAGTTGCTACCATCC 1440
QY 1441 TCAGTACCTCAACCAAGGCTCCTCTGTGTTCAACCGGGTCAACAGTGTGAATGCAATTGT 1500
DB 1441 TCAGTACCTCAACCAAGGCTCCTCTGTGTTCAACCGGGTCAACAGTGTGAATGCAATTGT 1500
QY 1501 CAAAGCAGGTCGAATGCTGAAAAATATAGAGGCTTGTGTGATGTTGGCCCAAGAA 1560
DB 1501 CAAAGCAGGTCGAATGCTGAAAAATATAGAGGCTTGTGTGATGTTGGCCCAAGAA 1560
QY 1561 ATTAACATGATTCGAGTACAGTGAAGCATGGGACCGAAGAAATATATGTTGGGGTGG 1620
DB 1561 ATTAACATGATTCGAGTACAGTGAAGCATGGGACCGAAGAAATATATGTTGGGGTGG 1620
QY 1621 CCTAGCTGATGTTTATGTTATTAATATATGTTGCTTATTAATTTTGCAGATGTTGAA 1680
DB 1621 CCTAGCTGATGTTTATGTTATTAATATATGTTGCTTATTAATTTTGCAGATGTTGAA 1680
QY 1681 TGCATCAGCTTCAATTAATGCTTTAGAGCGGGCATATCTGTTACTAGAAACATGAATG 1740
DB 1681 TGCATCAGCTTCAATTAATGCTTTAGAGCGGGCATATCTGTTACTAGAAACATGAATG 1740
QY 1741 AATGTATTAATTTTGTGT 1760
DB 1741 AATGTATTAATTTTGTGT 1760

RESULT 3
US-10-424-599-70167
; Sequence 70167, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 70167
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1
US-10-424-599-70167

Query Match 99.2%; Score 1745.8; DB 18; Length 1989;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1758; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 CTCTCTTATATCCCTTGTATATTCATTCATCTTAACTTTGTAATAATATATGTTCA 60
DB 164 CTCTCTTATATCCCTTGTATATTCATTCATCTTAACTTTGTAATAATATATGTTCA 223
QY 61 TCGAATATTTAAGGTGAGTGTCTTAATGTGAAGTACCGGAGCTGAGATTCACTCCG 120
DB 224 TCGAATATTTAAGGTGAGTGTCTTAATGTGAAGTACCGGAGCTGAGATTCACTCCG 283
QY 121 TGTACACATGAGAAACCCGGAATGTTGACAGAGAAAGGAAATGGCACTTCACTGTA 180
DB 284 TGTACACATGAGAAACCCGGAATGTTGACAGAGAAAGGAAATGGCACTTCACTGTA 343
QY 161 TGTCAAAACCCCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTTAATTAG 240
DB 344 TTGTCAAAACCCCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTTAATTAG 403
QY 241 GGGTAATGCTTGTGGGTT- GGGGTGAAACCAACGCGCTCAACCTTCAACCGGTGTATT 299
DB 404 GGGTAATGCTTGTGGGTTGGGGTGGAAACCAACGCGCTCAACCTTCAACCGGTGTATT 463
QY 300 GGTAAACCGAGAGGGAATTCATGGGCTAACAAAGCAAGATTCACAAAGCCAAATTACT 359
DB 464 GGTAAACAGAGAGGGAATTCATGGGCTAACAAAGCAAGATTCACAAAGCCAAATTACT 523
QY 360 GGCTCCCTCAACCAAGCTCAGCTATCCAGTGTGGGTCCTTCAGGAGAGGAAATCTAT 419
DB 524 GGCTCCCTCAACCAAGCTCAGCTATCCAGTGTGGGTCCTTCAGGAGAGGAAATCTAT 583
QY 420 GCCCATTCGAAGGCTGCTTCCATGTGTTAAACCTGACGACATGTGTGGGGATG 479
DB 584 GCCCATTCGAAGGCTGCTTCCATGTGTTAAACCTGACGACATGTGTGGGGATG 643
QY 480 GATATACGAAATATGAACCTGGGCTATGCGCATGGCCAGGCAAGGTGTTTAACTCAT 539
DB 644 GATATACGAAATATGAACCTGGGCTATGCGCATGGCCAGGCAAGGTGTTTAACTCAT 703
QY 540 TTGCAAGAGCAGTTGAGGCTTACATGGAATCATCTTCCACTCCCGGAATCTATGAC 599
DB 704 TTGCAAGAGCAGTTGAGGCTTACATGGAATCTATGGAATCTATGGAATCTATGAC 763
QY 600 CCGGATTTCAATGCTGCCAACAAGAGAGCGTGCACCAACGTCATCAAGGGCAACAAG 659
DB 764 CCGGATTTCAATGCTGCCAACAAGAGAGCGTGCACCAACGTCATCAAGGGCAACAAG 823
QY 660 CAAGAGCAGTTCAACAAATCATCAAAAGCATCAAGAGCTTTAAGGAAGCCCAAAAGTG 719
DB 824 CAAGAGCAGTTCAACAAATCATCAAAAGCATCAAGAGCTTTAAGGAAGCCCAAAAGTG 883
QY 720 GACAAGGTGTTGTAAGTGTGAGTGTGCAACAGAGAGGTACAGTAATTTGTTGGGAC 779
DB 884 GACAAGGTGTTGTAAGTGTGAGTGTGCAACAGAGAGGTACAGTAATTTGTTGGGAC 943
QY 780 GTTAAATGACACATGAGAAATCTTGGTGTGTGAGAGAAATGAGGCTGAGATTTCT 839
DB 944 GTTAAATGACACATGAGAAATCTTGGTGTGTGAGAGAAATGAGGCTGAGATTTCT 1003
QY 840 CTTTCAACCTTGTATGAGTGTGTGTGTTAAGAAATGTTCTTTTCAATTAATGGAAC 899
DB 1004 CTTTCAACCTTGTATGAGTGTGTGTGTTAAGAAATGTTCTTTTCAATTAATGGAAC 1063
QY 900 CCTCAAGACACTTTGTATCAGAGGCTGATGATCTTGCATGCGAGAGAACTTTGAT 959
DB 1064 CCTCAAGACACTTTGTATCAGAGGCTGATGATCTTGCATGCGAGAGAACTTTGAT 1123
QY 960 GGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTT 1019
DB 1019 GGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTT 1019
```

```

DB 1124 GGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTT 1183
QY 1020 GTGGGGGCTGGTATCAAGCAACATCTATAGTTCATCAACCATCTGGGAAACATGAT 1079
DB 1184 GTGGGGGCTGGTATCAAGCAACATCTATAGTTCATCAACCATCTGGGAAACATGAT 1243
QY 1080 GGTATGAATCTTGGGCTCCACAACTTTCGTTCCAAAGAAATCTCCAAAGCAAGTT 1139
DB 1244 GGTATGAATCTTGGGCTCCACAACTTTCGTTCCAAAGAAATCTCCAAAGCAAGTT 1303
QY 1140 GTTATGATATGCTACAGCAATGCCATCTCTATAGGCTGTGGTGAACATCCAGACAT 1199
DB 1304 GTTATGATATGCTACAGCAATGCCATCTCTATAGGCTGTGGTGAACATCCAGACAT 1363
QY 1200 GTTGTGTTATTAAGTATGTGCTTTCATGAGGAGACAGAGGAGCCATGATGATAC 1259
DB 1364 GTTGTGTTATTAAGTATGTGCTTTCATGAGGAGACAGAGGAGCCATGATGATAC 1423
QY 1260 ACTTCAGAGATATTCATGGGTGAAAGAGACCAATGTTTTCACACACATGCGAGAT 1319
DB 1424 ACTTCAGAGATATTCATGGGTGAAAGAGACCAATGTTTTCACACACATGCGAGAT 1483
QY 1320 TCCCTCTTACGCTGCTCTATATCTTGAATTTGGTCTTCTTGTGAGCTGACATGTA 1379
DB 1484 TCCCTCTTACGCTGCTCTATATCTTGAATTTGGTCTTCTTGTGAGCTGACATGTA 1543
QY 1380 ATCGAGTTTAAAGCTGAATGAGGAAATTCACATCTTCCACCCAGTTGCTACATC 1439
DB 1544 ATCGAGTTTAAAGCTGAATGAGGAAATTCACATCTTCCACCCAGTTGCTACATC 1603
QY 1440 CTCAGCTACCTCACCAAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGATGATG 1499
DB 1604 CTCAGCTACCTCACCAAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGATGATG 1663
QY 1500 TCAAGACGCTGCAATCTGGAATAACATAAGAGGCTTGTGTGATTTGGCCCCAGAG 1559
DB 1664 TCAAGACGCTGCAATCTGGAATAACATAAGAGGCTTGTGTGATTTGGCCCCAGAG 1723
QY 1560 AATAACATGATTTCTGAGTACAGAGTGAAGAGGAGCCAGAAATATATATGTTGGGTA 1619
DB 1724 AATAACATGATTTCTGAGTACAGAGTGAAGAGGAGCCAGAAATATATATGTTGGGTA 1783
QY 1620 GCTTAGCTGAATGTTTATGTTAATATATGTTTGTATTAATTTTGAAGTATTTGA 1679
DB 1784 GCTTAGCTGAATGTTTATGTTAATATATGTTTGTATTAATTTTGAAGTATTTGA 1843
QY 1680 ATGCATACGCTTCATTAATGCTTTAGACGGGCAATTTCTGTTTACTAGAACTGAT 1739
DB 1844 ATGCATACGCTTCATTAATGCTTTAGACGGGCAATTTCTGTTTACTAGAACTGAT 1903
QY 1740 GAATGTAGTATTAATTTGTGT 1760
DB 1904 GAATGTAGTATTAATTTGTGT 1924
```

```

RESULT 4
US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
```

NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 9
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-9

Query Match 87.1%; Score 1533; DB 14; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

54 ATGTCATGAGAAATTTAAGTTGAGTGTCTTAATGTGAAGTACCGAGATCGAGATT 113
1 ATGTCATGAGAAATTTAAGTTGAGTGTCTTAATGTGAAGTACCGAGATCGAGATT 60
114 CAGTCCGTGATACGAAACCAACGAACTTTTCAAGAAACAGAAATGACCTAT 173
61 CAGTCCGTGATACGAAACCAACGAACTTTTCAAGAAACAGAAATGACCTAT 120
174 CAGTCCGTGATACGAAACCAACGAACTTTTCAAGAAACAGAAATGACCTAT 233
121 CAGTCCGTGATACGAAACCAACGAACTTTTCAAGAAACAGAAATGACCTAT 180
234 AATTAGGGGTAATGCTTGTGGGTGGGTAAGAAACAGGCTCAACCTCAACGGTGT 293
181 AATTAGGGGTAATGCTTGTGGGTGGGTAAGAAACAGGCTCAACCTCAACGGTGT 240
294 GTTATTTGCTAACGAGGAGGATTTGATGAGGCTTCAAGAAACAGAAATGACCAAT 353
241 GTTATTTGCTAACGAGGAGGATTTGATGAGGCTTCAAGAAACAGAAATGACCAAT 300
354 TACTTTGGCTCCCTCAACCAAGCTTCAAGTTCGAGTTGGTCTTTCAAGGAGAGAA 413
301 TACTTTGGCTCCCTCAACCAAGCTTCAAGTTCGAGTTGGTCTTTCAAGGAGAGAA 360
414 ATCTATGCCCCAATTGAAGGCTGCTTCAATGTTAACTTGAAGCAATGTTGGTGG 473
361 ATCTATGCCCCAATTGAAGGCTGCTTCAATGTTAACTTGAAGCAATGTTGGTGG 420
474 GGATGGGATATCAGCAATGAACTGCTGATGCAATGAGGCAAGGCAAGGTTGAC 533
421 GGATGGGATATCAGCAATGAACTGCTGATGCAATGAGGCAAGGCAAGGTTGAC 480
534 ATCGATTTTCAAGAGAGTTGAGGCTTCAATGAAATCAATGCTTCACTCCCGGAATC 593
481 ATCGATTTTCAAGAGAGTTGAGGCTTCAATGAAATCAATGCTTCACTCCCGGAATC 540
594 TATGACCCGGATTTTCACTGCTGCAACCAAGAGAGGCTGCAACAGTCAATGAGGC 653
541 TATGACCCGGATTTTCACTGCTGCAACCAAGAGAGGCTGCAACAGTCAATGAGGC 600
654 ACAAGCAGAGAGGATTTCAACAAATCATCAAGATCAAGGCTTTAAAGAAAGCACC 713
601 ACAAGCAGAGAGGATTTCAACAAATCATCAAGATCAAGGCTTTAAAGAAAGCACC 660
714 AAAGTGACAGAGGTTTCACTGCTGCAACCAAGAGAGGCTTCACTCCCGGAATC 773
661 AAAGTGACAGAGGTTTCACTGCTGCAACCAAGAGAGGCTTCACTCCCGGAATC 720
774 GTGGGCTTAAATGACCAATGAGAAATCTTGGCTGCTGAGCAAGAAATGAGGCTGAG 833
721 GTGGGCTTAAATGACCAATGAGAAATCTTGGCTGCTGAGCAAGAAATGAGGCTGAG 780
834 ATTTCTCTTCCCACTTTGATGAGCAATGCTTGTGATGAAATGTTCTTTCAATTAAT 893
781 ATTTCTCTTCCCACTTTGATGAGCAATGCTTGTGATGAAATGTTCTTTCAATTAAT 840
894 GGAAGCCTTCAAGAACTTTTGTACAGAGGCTGATGATCTTCCAGAGGAAACACT 953
841 GGAAGCCTTCAAGAACTTTTGTACAGAGGCTGATGATCTTCCAGAGGAAACACT 900
954 TTGATTTGGTGAAGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGGTGAAT 1013

DB 901 TTGATTTGGTGAAGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
1014 TTCTTTGGGGGCTGATATCAAGCCAACTATATGTCATTAACCAATCTGGAAAC 1073
DB 961 TTCTTTGGGGGCTGATATCAAGCCAACTATATGTCATTAACCAATCTGGAAAC 1020
1074 AATGATGATGAAATCTTTGGGCTCCAAACTTTCCGTTCCAAAGAAATCTCCAAAGAC 1133
1021 AATGATGATGAAATCTTTGGGCTCCAAACTTTCCGTTCCAAAGAAATCTCCAAAGAC 1080
1134 AACGTTGTGATGATATGATGATCAAGCAATGCAATCTTATGAGGCTGTAACATCA 1193
1081 AACGTTGTGATGATATGATGATCAAGCAATGCAATCTTATGAGGCTGTAACATCA 1140
1194 GACCAATGTTGTTATTAATGATGATGCTTCAAGGAGGAGCAAGAGCAATGAT 1253
1141 GACCAATGTTGTTATTAATGATGATGCTTCAAGGAGGAGCAAGAGCAATGAT 1200
1254 GAGTACATTTCAAGATATTTCAATGAGTGAAGAGCAATGTTTGAACAACATGCT 1313
1201 GAGTACATTTCAAGATATTTCAATGAGTGAAGAGCAATGTTTGAACAACATGCT 1260
1314 GAGATTTCCCTCTTGAAGCTGCTCTTATATCTTGAAGCTTCTTCTGAGCTACG 1373
1261 GAGATTTCCCTCTTGAAGCTGCTCTTATATCTTGAAGCTTCTTCTGAGCTACG 1320
1374 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTTCACTCAATCCAGTTGCT 1433
1321 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTTCACTCAATCCAGTTGCT 1380
1434 ACCATCTTCAAGTACTTCAACCAAGGCTCTCTGTTTCCACGGGTAACCAATGATGAT 1493
1381 ACCATCTTCAAGTACTTCAACCAAGGCTCTCTGTTTCCACGGGTAACCAATGATGAT 1440
1494 GCATTTGCAAGCAGGTCGATGCTGAGAAACATATGAGGCTTGTGTTGATGGCC 1553
1441 GCATTTGCAAGCAGGTCGATGCTGAGAAACATATGAGGCTTGTGTTGATGGCC 1500
1554 CCAGAAATTAACATGATTTCTGAGTACAAAGTGA 1586
1501 CCAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 5
US-10-718-952-9
Sequence 9, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAPEINOS
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 9
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-718-952-9
Query Match 87.1%; Score 1533; DB 19; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

54 ATGTTTCATCGAGAAATTTTAAAGGTGAGTGTCTCTAAATGGAATGACACCGAGCTGAGATT 113
1 ATGTTTCATCGAGAAATTTTAAAGGTGAGTGTCTCTAAATGGAATGACACCGAGCTGAGATT 60
114 CAGTCGCGTGTACAACTAGCAAAACAACGAACTTGTTCAGAGAAACAGAAATGGACCTAT 173
61 CAGTCGCGTGTACAACTAGCAAAACAACGAACTTGTTCAGAGAAACAGAAATGGACCTAT 120
174 CAGTCGCGTGTACAACTAGCAAAACAACGAACTTGTTCAGAGAAACAGAAATGGACCTAT 233
121 CAGTCGCGTGTACAACTAGCAAAACAACGAACTTGTTCAGAGAAACAGAAATGGACCTAT 180
234 AAATTAAGGGGTAAATGTTGGGTGGGGTGGAAACAACGCTCAACCTCAACCGGTGAT 293
181 AAATTAAGGGGTAAATGTTGGGTGGGGTGGAAACAACGCTCAACCTCAACCGGTGAT 240
234 GTTATTTGCTAACCGAGAGGCAATTCATGGGCTACAAAGAGCAAGATTCACCAAGCAAT 353
241 GTTATTTGCTAACCGAGAGGCAATTCATGGGCTACAAAGAGCAAGATTCACCAAGCAAT 300
354 TACTTTGGCTCCCTCAACCAAGCTCAAGTATCCGAGTTGGGTCTTCCAGAGGAGAGAA 413
301 TACTTTGGCTCCCTCAACCAAGCTCAAGTATCCGAGTTGGGTCTTCCAGAGGAGAGAA 360
414 ATCTATGCCCAATTCAGAGAGCTGCTTCCAAATGTTAACCTGACGACATTTGTTGGG 473
361 ATCTATGCCCAATTCAGAGAGCTGCTTCCAAATGTTAACCTGACGACATTTGTTGGG 420
474 GGATGGGATATCAGCAACATGAACCTGCTGATGATGCGATGGCCAGGCAAAAGGTGTTGAC 533
421 GGATGGGATATCAGCAACATGAACCTGCTGATGATGCGATGGCCAGGCAAAAGGTGTTGAC 480
534 ATCGATTTGCAAGAGAGCTTGAAGGCTTACATGGAATCAGTCTTCCATCTCCCGGAATC 593
481 ATCGATTTGCAAGAGAGCTTGAAGGCTTACATGGAATCAGTCTTCCATCTCCCGGAATC 540
594 TATGACCCGGATTTTCAATGCTGCAACCAAGAGAGCGTCCAAACAGCTCATCAAGGAC 653
541 TATGACCCGGATTTTCAATGCTGCAACCAAGAGAGCGTCCAAACAGCTCATCAAGGAC 600
654 ACNAAAGCAGAGCAAGTTCAACAAATCATCAAGATCAGAGCGCTTTAAGAGACCAAC 713
601 ACNAAAGCAGAGCAAGTTCAACAAATCATCAAGATCAGAGCGCTTTAAGAGACCAAC 660
714 AAAGTGGACAAAGGTGTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 773
661 AAAGTGGACAAAGGTGTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 720
774 GTGGGCTTAATGACACATGAGAAATCTCTGGCTGCTGTGGAACAGAAATGAGCTGAG 833
721 GTGGGCTTAATGACACATGAGAAATCTCTGGCTGCTGTGGAACAGAAATGAGCTGAG 780
834 ATTTCTCTTCCACCTTGTATGCAATGCTTGTATGGAATGTTCTTTCATTAAT 893
781 ATTTCTCTTCCACCTTGTATGCAATGCTTGTATGGAATGTTCTTTCATTAAT 840
894 GGAAGCCCTCAAGAACTTTTGTACAGAGGCTGATTTGCTTCCAGAGAAACAT 953
841 GGAAGCCCTCAAGAACTTTTGTACAGAGGCTGATTTGCTTCCAGAGAAACAT 900
954 TTGATTTGGGAGATGACTTCAAGAGTGGTCAAGACCAAAATGAAATCTGTTGGTGAAT 1013
901 TTGATTTGGGAGATGACTTCAAGAGTGGTCAAGACCAAAATGAAATCTGTTGGTGAAT 960
1014 TTCTTTGGGGGCTGTATCAAGCCACATCTATATGCAATTAACCAATCTGGGAAAC 1073
961 TTCTTTGGGGGCTGTATCAAGCCACATCTATATGCAATTAACCAATCTGGGAAAC 1020
1074 AATGATGATGATATTTTGGGCTCAAGAACTTTCCGTTCCAGAGAAATCTTCAAGAC 1133
1021 AATGATGATGATATTTTGGGCTCAAGAACTTTCCGTTCCAGAGAAATCTTCAAGAC 1080

QY 1134 AACGTTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
DB 1081 AACGTTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1194 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
DB 1141 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1254 GAGTACATTCAGAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
DB 1201 GAGTACATTCAGAGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1314 GAGATTCCTCTTATGCTGCTCTTATTAATGATGATGATGATGATGATGATGATGAT 1373
DB 1261 GAGATTCCTCTTATGCTGCTCTTATTAATGATGATGATGATGATGATGATGATGAT 1320
QY 1374 ACTAGATTCAGATTTAAAGCTGAAATTAAGGAAATTCACATCAATTCACCAAGTTGCT 1433
DB 1321 ACTAGATTCAGATTTAAAGCTGAAATTAAGGAAATTCACATCAATTCACCAAGTTGCT 1380
QY 1434 ACCATCTCAGCTACCTGACCAAGGCTCTCTGTTCCAGCGGGTACACAGTGTGAT 1493
DB 1381 ACCATCTCAGCTACCTGACCAAGGCTCTCTGTTCCAGCGGGTACACAGTGTGAT 1440
QY 1494 GCATTGTCAAGCAAGCGTGCATATGCTGGAATAATATAGAGGCTTGTGTTGATGATG 1553
DB 1441 GCATTGTCAAGCAAGCGTGCATATGCTGGAATAATATAGAGGCTTGTGTTGATGATG 1500
QY 1554 CCAGGAATTAACATGATTTCTGAGTACAAAGTGA 1586
DB 1501 CCAGGAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 6

US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-5

Query Match 87.0%; Score 1531.4; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 ATGTTTCATCGAGAAATTTTAAAGGTGAGTGTCTCTAAATGGAATGACACCGAGCTGAGATT 113
DB 1 ATGTTTCATCGAGAAATTTTAAAGGTGAGTGTCTCTAAATGGAATGACACCGAGCTGAGATT 60
QY 114 CAGTCGCGTGTACAACTAGCAAAACAACGAACTTGTTCAGAGAAACAGAAATGGACCTAT 173
DB 61 CAGTCGCGTGTACAACTAGCAAAACAACGAACTTGTTCAGAGAAACAGAAATGGACCTAT 120
QY 174 CAGTCGCGTGTACAACTAGCAAAACAACGAACTTGTTCAGAGAAACAGAAATGGACCTAT 233

Db 121 CAGTGATGTCACAAACCAATCTGTCAAAATGAAATTTAAACCAACATCCATGTTCT 180
Qy 234 AAATTAGGGGTATATGCTTGGGTTGGGGTGAATAAGGCTCAACCTTCAACCGGTGGT 293
Db 181 AAATTAGGGGTATATGCTTGGGTTGGGGTGAATAAGGCTCAACCTTCAACCGGTGGT 240
Qy 294 GTTATTTGCTAACGAGAGGCAATTTATGAGGCTTACAAAGGACAAAGATTTCAACAGCCAT 353
Db 241 GTTATTTGCTAACGAGAGGCAATTTATGAGGCTTACAAAGGACAAAGATTTCAACAGCCAT 300
Qy 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGTCTTCCAGGGAGAGAA 413
Db 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGTCTTCCAGGGAGAGAA 360
Qy 414 ATCTATGCCCCATTCAAGAGCCGTCTTCAATGCTTAACTTGAAGACATTTGTTGGG 473
Db 361 ATCTATGCCCCATTCAAGAGCCGTCTTCAATGCTTAACTTGAAGACATTTGTTGGG 420
Qy 474 GGATGGATATACGACAACTGAACCTGGCTGATGCGATGGCCAGGGCAAGGTTTGAAC 533
Db 421 GGATGGATATACGACAACTGAACCTGGCTGATGCGATGGCCAGGGCAAGGTTTGAAC 480
Qy 534 ATCGATTTGACAGAGCAGTTGAGGCTTACATGGAATCCATGCTTCCCTCCGGAATC 593
Db 481 ATCGATTTGACAGAGCAGTTGAGGCTTACATGGAATCCATGCTTCCCTCCGGAATC 540
Qy 594 TATGACCCCGATTTTCTTCTGCTCAACGAGAGAGCGTCCAAACAGCTCATCAAGGCG 653
Db 541 TATGACCCCGATTTTCTTCTGCTCAACGAGAGAGCGTCCAAACAGCTCATCAAGGCG 600
Qy 654 ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGAGCATCAAGGCGTTTAAAGAGCCAC 713
Db 601 ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGAGCATCAAGGCGTTTAAAGAGCCAC 660
Qy 714 AAAGTGAACAAGTGGTGTGCTGCTGCAACAGAGAGGATCAGTATTTGGT 773
Db 661 AAAGTGAACAAGTGGTGTGCTGCTGCAACAGAGAGGATCAGTATTTGGT 720
Qy 774 GTGGGCTTTAATGACACCACTGAGAACTCTTGGCTGCTGCTGAGACAAATAGCTGAG 833
Db 721 GTGGGCTTTAATGACACCACTGAGAACTCTTGGCTGCTGCTGAGACAAATAGCTGAG 780
Qy 834 ATTTCTCTTCCACCTGTGATGCAATGCTTGTGTTAGGAAATGCTTCTTCAATAT 893
Db 781 ATTTCTCTTCCACCTGTGATGCAATGCTTGTGTTAGGAAATGCTTCTTCAATAT 840
Qy 894 GGAAGCCCTCAGACACTTTGTACAGAGGCTGATTTGATCTTGCATCGAGAACT 953
Db 841 GGAAGCCCTCAGACACTTTGTGTACAGAGGCTGATTTGATCTTGCATCGAGAACT 900
Qy 954 TTGATTTGGTGAATGACTTCAAGAGTGTGACAAATGAATCTGTGTTGGTAT 1013
Db 901 TTGATTTGGTGAATGACTTCAAGAGTGTGACAAATGAATCTGTGTTGGTAT 960
Qy 1014 TTCTCTTGGGGGGCTGTATCAAGGCAACATCTATATAGCTTACACATCTTGGGAAAC 1073
Db 961 TTCTCTTGGGGGGCTGTATCAAGGCAACATCTATATAGCTTACACATCTTGGGAAAC 1020
Qy 1074 AATGATGATGAATCTTTCGCTCCAAACCTTTCGCTCCAAAGAAATCTCAAGAGC 1133
Db 1021 AATGATGATGAATCTTTCGCTCCAAACCTTTCGCTCCAAAGAAATCTTCAAGAGC 1080
Qy 1134 AAAGTTGATGATATGCTCAACAGCAATGCTATCTTATAGCTTGTGATCAATCA 1193
Db 1081 AAAGTTGATGATATGCTCAACAGCAATGCTATCTTATAGCTTGTGATCAATCA 1140
Qy 1194 GACCAATGTTGTTATTAAGATATGCTTACGTAAGGGAACAGCAAGAGCCATGAT 1253
Db 1141 GACCAATGTTGTTATTAAGATATGCTTACGTAAGGGAACAGCAAGAGCCATGAT 1200
Qy 1254 GAGTACACTTCAAGATATTTCAATGAGTGAAGAGACCAATGTTTTCACAAACATGC 1313
Db 1201 GAGTACACTTCAAGATATTTCAATGAGTGAAGAGACCAATGTTTTCACAAACATGC 1260

Qy 1314 GAGATTCCTCTAGCTGCTCTATATATCTTGAATTTGGTCTTCTTGTGAGCTACG 1373
Db 1261 GAGATTCCTCTAGCTGCTCTATATATCTTGAATTTGGTCTTCTTGTGAGCTACG 1320
Qy 1374 ACTAGATGAGTTTAAAGCTGAAATGAGGGAATTTCCATCTATTTCAACCAAGTTGCT 1433
Db 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGGAATTTCCATCTATTTCAACCAAGTTGCT 1380
Qy 1434 ACCATCTCAGCTTACCTCAACAGGCTCTGTTTCCACCGGGTACACAGTGTGAT 1493
Db 1381 ACCATCTCAGCTTACCTCAACAGGCTCTGTTTCCACCGGGTACACAGTGTGAT 1440
Qy 1494 GATTTGCTAACAGCGCTGCTGATGCTGAGAAACATATATAGGCTTGTGATGGCC 1553
Db 1441 GATTTGCTAACAGCGCTGCTGATGCTGAGAAACATATATAGGCTTGTGATGGCC 1500
Qy 1554 CCAGAAATATACATGATTTCTGAGTACAAAGTGA 1586
Db 1501 CCAGAAATATACATGATTTCTGAGTACAAAGTGA 1533

RESULT 7
US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-5

Query Match 87.0%; Score 1531.4; DB 19; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 54 ATGTCATCGGAATTTTAAAGTGTGATGCTCTATATGAGTATACCGAGCTGAGATT 113
Db 1 ATGTCATCGGAATTTTAAAGTGTGATGCTCTATATGAGTATACCGAGCTGAGATT 60
Qy 114 CAGTCCGTGTAACAATAAGCAACCAAGCACTTGTTCAGAGAACAGAAATGGCACTAT 173
Db 61 CAGTCCGTGTAACAATAAGCAACCAAGCACTTGTTCAGAGAACAGAAATGGCACTAT 120
Qy 174 CAGTGGATTTGCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATTTCT 233
Db 121 CAGTGGATTTGCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATTTCT 180
Qy 234 AAATTAGGGGTATATGCTTGGGTTGGGGTGAATAAGGCTCAACCTTCAACCGGTGGT 293
Db 181 AAATTAGGGGTATATGCTTGGGTTGGGGTGAATAAGGCTCAACCTTCAACCGGTGGT 240
Qy 294 GTTATTTGCTAACGAGAGGCAATTTATGAGGCTTACAAAGGACAAAGATTTCAACAGCCAT 353
Db 241 GTTATTTGCTAACGAGAGGCAATTTATGAGGCTTACAAAGGACAAAGATTTCAACAGCCAT 300
Qy 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGTCTTCCAGGGAGAGAA 413


```
Db 301 TACTTGGCTCCCTCAACCCAAAGCTCAGCTATCCAGTGGGTCTTCCAGGAGAGAA 360
Qy 414 ATCTATGCCCATTCGAAGAGCTGCTTCCATGTGTTAAACCTGACGACATTTGTTGGG 473
Db 361 ATCTATGCCCATTCGAAGAGCTGCTTCCATGTGTTAAACCTGACGACATTTGTTGGG 420
Qy 474 GGATGGGATATCAGCAACATGAACCTGGCTGATGCGCATGGGCAAAAGTGTGTTAC 533
Db 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCGCATGGGCAAAAGTGTGTTAC 480
Qy 534 ATCGATTTGACAGACAGCTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGATC 593
Db 481 ATCGATTTGACAGACAGCTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGATC 540
Qy 594 TATGACCCGGATTTATGCTGCTCCAAACGAAGAGCGTCCAAACAGTCAATCAAGGAC 653
Db 541 TATGACCCGGATTTATGCTGCTCCAAACGAAGAGCGTCCAAACAGTCAATCAAGGAC 600
Qy 654 ACAAGCAAGACAGCAAGTTCACAAATCATCAAAAGCATCAAGCGTTTAAAGAACCAAC 713
Db 601 ACAAGCAAGACAGCAAGTTCACAAATCATCAAAAGCATCAAGCGTTTAAAGAACCAAC 660
Qy 714 AAAGTGAACAAGTGTGTTGATCTGTGACCTGCCAACACAGAGGTACATTAATTGGTT 773
Db 661 AAAGTGAACAAGTGTGTTGATCTGTGACCTGCCAACACAGAGGTACATTAATTGGTT 720
Qy 774 GTGGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGAGACAGAAATGAGGCTAG 833
Db 721 GTGGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGAGACAGAAATGAGGCTAG 780
Qy 834 ATTTCTCTTCCACCTTGTATGCAATGCTTGTGTTATGAAATGTTCTTTCTTAAT 893
Db 781 ATTTCTCTTCCACCTTGTATGCAATGCTTGTGTTATGAAATGTTCTTTCTTAAT 840
Qy 894 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAACT 953
Db 841 GGAAGCCCTCAGAACATTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAACT 900
Qy 954 TTGATTTGGTGAATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGAT 1013
Db 901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGAT 960
Qy 1014 TTCTTGTGGGGCTGTATCAAGCCAAATCTATAGTACATCAACCATCTGGAAAC 1073
Db 961 TTCTTGTGGGGCTGTATCAAGCCAAATCTATAGTACATCAACCATCTGGAAAC 1020
Qy 1074 AATGATGATGAATCTTTCGGCTCACAACATTTCCGTTCCAAAGAAATCTCCAAGAC 1133
Db 1021 AATGATGATGAATCTTTCGGCTCACAACATTTCCGTTCCAAAGAAATCTCCAAGAC 1080
Qy 1134 AACGTTTGTATGATATGATCAACAGCAATGCAATCTCTATAGACCTGTGTACATCA 1193
Db 1081 AACGTTTGTATGATATGATCAACAGCAATGCAATCTCTATAGACCTGTGTACATCA 1140
Qy 1194 GACCAATGTTGTTATTAAGTATGTGCTTACGTAAGGGAACAGAAAGAACCATGAT 1253
Db 1141 GACCAATGTTGTTATTAAGTATGTGCTTACGTAAGGGAACAGAAATGACCATGAT 1200
Qy 1254 GAGTACATTCAGAGATATTCATGSGTGAAGAGACCATTTGTTGCAACACATGAC 1313
Db 1201 GAGTACATTCAGAGATATTCATGSGTGAAGAGACCATTTGTTGCAACACATGAC 1260
Qy 1314 GAGGATTCCTCTTACGCTCTCTATTAATCTTGAATGCTGCTCTTGTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTTACGCTCTCTATTAATCTTGAATGCTGCTCTTGTGAGCTCAGC 1320
Qy 1374 ACTAGAAATGAAATTAAGCTGAAATGAGGAAATTCCTCACTTCCACCACTTGTCT 1433
Db 1321 ACTAGAAATGAAATTAAGCTGAAATGAGGAAATTCCTCACTTCCACCACTTGTCT 1380
Qy 1434 ACCATCTCAGTACCTCACAAGGCTCTCTGTGTTCCACCGGGTACCAAGTGTGAT 1493
Db 1493 ACCATCTCAGTACCTCACAAGGCTCTCTGTGTTCCACCGGGTACCAAGTGTGAT 1450
```

```
Db 1381 AACATCTCAGCTACCTCACCACCAAGGCTCTCTGTTCCACCGGGTACACCAAGTGTGAT 1440
Qy 1494 GCATTTGCAAAAGCAGCGTGCATATCTGAAAAACATAATGAGGCTTGTGATGAGCC 1553
Db 1441 GCATTTGCAAAAGCAGCGTGCATATCTGAAAAACATAATGAGGCTTGTGATGAGCC 1500
Qy 1554 CCAAGAAATTAACATGATTTCTGAGTACAGTGA 1586
Db 1501 CCAGAGAAATTAACATGATTTCTGAGTACAGTGA 1533

RESULT 8
US-10-025-003-13
; Sequence 13, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-13

Query Match 83.3%; Score 1465.8; DB 14; Length 1533;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 54 ATGTTTCATCGAATTTTAAGTGTGATGCTCTATGTAAGTACACCGAGATCGAGATT 113
Db 1 ATGTTTCATCGAATTTTAAGTGTGATGCTCTATGTAAGTACACCGAGATCGAGATT 60
Qy 114 CAGTCCGTTGTAACATACGAACCAACCACTTGTTCAGAGAACAGAAATGCACTAT 173
Db 61 CAGTCCGTTGTAACATACGAACCAACCACTTGTTCAGAGAACAGAAATGCACTAT 120
Qy 174 CAGTGGATTTGCAAAACCAATCTGTCAAAATAGAAATTTAAACCAACATCATGTTCT 233
Db 121 CAGTGGATTTGCAAAACCAATCTGTCAAAATAGAAATTTAAACCAACATCATGTTCT 180
Qy 234 AAATTAAGGGGTAATGCTTGTGGGTGGGTTGGGTAACAACCGCTCAACCTTACCGGTGT 293
Db 181 AAATTAAGGGGTAATGCTTGTGGGTGGGTTGGGTAACAACCGCTCAACCTTACCGGTGT 240
Qy 294 GTTATGTTAAACGAGAGGCAATTTCAATGAGCTCAAAAGACAGAAATTAACAAGCAAT 353
Db 241 GTTATGTTAAACGAGAGGCAATTTCAATGAGCTCAAAAGACAGAAATTAACAAGCAAT 300
Qy 354 TACTTTGGCTCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 413
Db 301 TACTTTGGCTCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 360
Qy 414 ATCTATGCCCATTCGAAGAGCTGCTTCCATGTGTTAAACCTGACGACATTTGTTGGG 473
Db 361 ATCTATGCCCATTCGAAGAGCTGCTTCCATGTGTTAAACCTGACGACATTTGTTGGG 420
Qy 474 GGATGGGATATCAGCAACATGAACCTGGCTGATGCGCATGGGCAAAAGTGTGTTAC 533
Db 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCGCATGGGCAAAAGTGTGTTAC 480
```

Oy 534 ATCGATTTCAGAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAAATC 593
 Db 481 ATCGATTTCAGAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAAATC 540
 Oy 594 TATGACCCGAGTTTCTGCTGCAACCAAGAGAGCTGCAACCAAGCTCATCAAGGCG 653
 Db 541 TACGACCCGAGTTTCTGCTGCAACCAAGAGAGCTGCAACCAAGCTCATCAAGGCG 600
 Oy 654 ACAAGCAAGAGAGTTCAACCAATCATCAAGAGATCAAGGCGTTTAAAGAGGCAAC 713
 Db 601 ACAAGCAAGAGAGTTCAAGCAATCATCAAGAGATCAAGGCGTTTAAAGAGGCAAC 660
 Oy 714 AAAGTGACAGAGTGTGTACTGTGAGCTCCCAACAGAGAGTAAATTTGGTT 773
 Db 661 AAAGTGACAGAGTGTGTACTGTGAGCTCCCAACAGAGAGTAAATTTGGTT 720
 Oy 774 GTGGGCTTAAATGACCAATGAGAAATCTTGTGCTGCTGTGCAAGAAATGAGGCTGAG 833
 Db 721 GTAGGCTTAAATGACCAATGAGAAATCTTGTGCTGCTGTGCAAGAAATGAGGCTGAG 780
 Oy 834 ATTTCTCTTCCACCTGTATGCAATGCTGTGTATGGAATAATGCTCTTCAATAT 893
 Db 781 ATTTCTCTTCCACCTGTATGCAATGCTGTGTATGGAATAATGCTCTTCAATAT 840
 Oy 894 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCATCGCAGAAACAT 953
 Db 841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCATCGCAGAAACAT 900
 Oy 954 TTGATTTGGTGAAGTACTTCAAGAGTGTGCAAGCAAAATGAAATCTGTGTTGTTAT 1013
 Db 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGCAAGCAAAATGAAATCTGTGTTGTTAT 960
 Oy 1014 TTCCTTGGGGGCTGTATCAAGGCAATCATGTATGATTAACAACATCGGGAAC 1073
 Db 961 TTCCTTGGGGGCTGTATCAAGGCAATCATGTATGATTAACAACATCGGGAAC 1020
 Oy 1074 AATGATGTATGATCTTTCGCTCCCAAACTTTCGTTCCAAAGAAATCTCCAAAGAC 1133
 Db 1021 AATGATGTATGATCTTTCGCTCCCAAACTTTCGTTCCAAAGAAATCTCCAAAGAC 1080
 Oy 1134 AAGTTTGTATGATATGCTCAACAGCAATGCTTATGAGCTGTGTAACATCCA 1193
 Db 1081 AAGTTTGTATGATATGCTCAACAGCAATGCTTATGAGCTGTGTAACATCCA 1140
 Oy 1194 GACCAATGTTGTTATTAATGATGCTTACATGAGGGAACAGCAAGAGGCTATGAT 1253
 Db 1141 GACCAATGTTGTTATTAATGATGCTTACATGAGGGAATGCAAGAGGCTATGAT 1200
 Oy 1254 GAGTACACTTCAAGATATTTCAATGAGTGAAGAGCAATGTTTGTGCAACACATGC 1313
 Db 1201 GAGTACACTTCAAGATATTTCAATGAGTGAAGAGCAATGTTTGTGCAACACATGC 1260
 Oy 1314 GAGATTCCTCTTACAGCTCTCTTATTAATCTTGAATGCTGTCTTGTGCTGAGCTAC 1373
 Db 1261 GAGATTCCTCTTACAGCTCTCTTATTAATCTTGAATGCTGTCTTGTGCTGAGCTAC 1320
 Oy 1374 ACTAATATGAGTATTAAGCTGAAATGAGGAAATTCATCTATCCACCAAGTTGCT 1433
 Db 1321 ACTAATATGAGTATTAAGCTGAAATGAGGAAATTCATCTATCCACCAAGTTGCT 1380
 Oy 1434 ACCATCTCAGCTACCTCAACAGGCTCTCTGTTCCACCGGATACACAGTGTGTAAT 1493
 Db 1381 ACCATCTCAGCTACCTCAACAGGCTCTCTGTTCCACCGGATACACAGTGTGTAAT 1440
 Oy 1494 GCATTGTCAAGACGCTGCAATGCTGAAACATTAATGAGGCTGTGTTGATTTGGCC 1553
 Db 1441 GCATTGTCAAGACGCTGCAATGCTGAAACATTAATGAGGCTGTGTTGATTTGGCC 1500
 Oy 1554 CCAGAGATTAACATGATCTCGAGTCAAGTGA 1586
 Db 1501 CCAGAGATTAACATGATCTCGAGTCAAGTGA 1533

RESULT 9
 US-10-025-003-15
 ; Sequence 15, Application US/10025003
 ; Publication No. US20030074685A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hiltz, William
 ; APPLICANT: Sebastian, Scott
 ; APPLICANT: Grace, John
 ; APPLICANT: Streitz, Leon
 ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
 ; FILE REFERENCE: BB-1077-C
 ; CURRENT APPLICATION NUMBER: US/10/025,003
 ; PRIOR FILING DATE: 2002-05-07
 ; PRIOR APPLICATION NUMBER: 08/835,751
 ; PRIOR FILING DATE: APRIL 8, 1997
 ; PRIOR APPLICATION NUMBER: PCT/US98/06822
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO. 15
 ; TYPE: DNA
 ; LENGTH: 1533
 ; ORGANISM: Glycine max
 ; US-10-025-003-15
 Query Match 83.3%; Score 1465.8; DB 14; Length 1533;
 Best Local Similarity 97.3%; Pred. No. 0;
 Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 Oy 54 ATGTCATGAGAAATTTTAAAGTGTGCTTCAATGTAAGTACACGAGACTGAGATT 113
 Db 1 ATGTCATGAGAAATTTTAAAGTGTGCTTCAATGTAAGTACACGAGACTGAGATT 60
 Oy 114 CAGTCGTGTACACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGCACTAT 173
 Db 61 CAGTCGTGTACACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGCACTAT 120
 Oy 174 CAGTGGATGTCAAAACCAATCTGTCAATATGCAATTTAAACCAACATCTGTTCT 233
 Db 121 CAGTGGATGTCAAAACCAATCTGTCAATATGCAATTTAAACCAACATCTGTTCT 180
 Oy 234 AAATTAAGGGGTAATGCTTGTGGGTGGAACAACGGCTCAACCTGACCGGTGAT 293
 Db 181 AAATTAAGGGGTAATGCTTGTGGGTGGAACAACGGCTCAACCTGACCGGTGAT 240
 Oy 294 GTTATTTGCTAACGAGAGGCAATTCATGAGGCTAACAAAGAACAGATTCAACAGCAAT 353
 Db 241 GTTATTTGCTAACGAGAGGCAATTCATGAGGCTAACAAAGAACAGATTCAACAGCAAT 300
 Oy 354 TACTTTGGCTCCCTCAACCAAGCTTCAAGTATCCAGATTGGGTCTTCCAGGAGAGAA 413
 Db 301 TACTTTGGCTCCCTCAACCAAGCTTCAAGTATCCAGATTGGGTCTTCCAGGAGAGAA 360
 Oy 414 ATCTATGCCCCATTCAGAGGCTGCTTCAATGTTAAACCTCGAAGCAATGTGTTGGG 473
 Db 361 ATCTATGCCCCATTCAGAGGCTGCTTCAATGTTAAACCTCGAAGCAATGTGTTGGG 420
 Oy 474 GGATGGATATTCAGACATGAACTGCTGATGCTATGCTGCAAGGAGCAAGAGTGTGAC 533
 Db 421 GGATGGATATTCAGACATGAACTGCTGATGCTATGCTGCAAGGAGCAAGAGTGTGAC 480
 Oy 534 ATGATTTGCAAGACAGTTGAGGCTTATCATGAAATCATGCTTCCACTCCCGAATC 593
 Db 481 ATGATTTGCAAGACAGTTGAGGCTTATCATGAAATCATGCTTCCACTCCCGAATC 540
 Oy 594 TATGACCCGGAATTTATGCTGCAACCAAGAGAGGCTGCAACAGTCAATCAAGGCG 653
 Db 541 TATGACCCGGAATTTATGCTGCAACCAAGAGAGGCTGCAACAGTCAATCAAGGCG 600
 Oy 654 ACAAGCAAGAGCAAGTTCAACCAATCATCAAGAGATCAAGGCTTAAAGAGGCAAC 713
 Db 601 ACAAGCAAGAGCAAGTTCAACCAATCATCAAGAGATCAAGGCTTAAAGAGGCAAC 660

QY 714 AAAGTGAACAAGTGGTGTGACTGTGAGCTGGCCAAACAGAGAGGTAGTAATTTGGT 773
DB 661 AAAGTGAACAAGTGGTGTGACTGTGAGCTGGCCAAACAGAGAGGTAGTAATTTGGT 720
QY 774 GTGGGCTTAATGACACCAATGAGAAATCTTTGGTCTGTGTGACAGAAATGAGGCTGAG 833
DB 721 GTAGGCTTAATGACACCAATGAGAAATCTTTGGTCTGTGTGACAGAAATGAGGCTGAG 780
QY 834 ATTTCTCTTCCACCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 893
DB 781 ATTTCTCTTCCACCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 840
QY 894 GGAAGCCCTCAGAAACACTTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAAACT 953
DB 841 GGAAGCCCTCAGAAACACTTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAAACT 900
QY 954 TTGATTTGGTGAAGATGACTTCAAGAGTGTGTGACACCAAAATGAAATCTGTGTGAT 1013
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTGTGACACCAAAATGAAATCTGTGTGAT 960
QY 1014 TTCTTTGGGGGGCTGTGTGACAGCAATCTATATGTCAGTTACAACTCTGGGAAAC 1073
DB 961 TTTCTTTGGGGGGCTGTGTGACAGCAATCTATATGTCAGTTACAACTCTGGGAAAC 1020
QY 1074 AATGATGGTATGATCTTTCCGCTCCAAACCTTCCGTTCCAGAGAAATCTCAAGAGC 1133
DB 1021 AATGATGGTATGATCTTTCCGCTCCAAACCTTCCGTTCCAGAGAAATCTCAAGAGC 1080
QY 1134 AACGTGTGTGATGATATGTCACACCAATGTCATCTTATGAGCTGTGTGAAATCCA 1193
DB 1081 AACGTGTGTGATGATATGTCACACCAATGTCATCTTATGAGCTGTGTGAAATCCA 1140
QY 1194 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTTAGGGGGAACAGCAAGAGCCATGAT 1253
DB 1141 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTTAGGGGGAATGACAAAGAGCCATGAT 1200
QY 1254 GAGTACACTTCAAGATATTCATGGGTGGAAGAGACCATTTGTTGACAAACATGSC 1313
DB 1201 GAGTACACTTCAAGATATTCATGGGTGGAAGAGACCATTTGTTGACAAACATGSC 1260
QY 1314 GAGGATTCCTCTTACGCTCTCTATTTATCTTGAATGTTGCTCTTCTGTGAGCTCAGC 1373
DB 1261 GAGGATTCCTCTTACGCTCTCTATTTATCTTGAATGTTGCTCTTCTGTGAGCTCAGC 1320
QY 1374 ACTAGATGAGTTTAAAGTGAAGGGAATTTCCATCTCATTTCCACCAGTTGCT 1433
DB 1321 ACTAGATGAGTTTAAAGTGAAGGGAATTTCCATCTCATTTCCACCAGTTGCT 1380
QY 1434 ACCATCTCAGCTACCTCAAGAGCTCTGTGTTCCACCGGTTACCAAGTGTGAT 1493
DB 1381 ACCATCTCAGCTATCTGACCAAGGCTCTGTGTTCCACCGGTTACCAAGTGTGAT 1440
QY 1494 GCATTGTCAAGCAGCGTCAATGCTGAAAAATTAATGAGGCTGTGTGATTTGAGC 1553
DB 1441 GCATTGTCAAGCAGCGTCAATGCTGAAAAATTAATGAGGCTGTGTGATTTGAGC 1500
QY 1554 CCAAGAGATTAATGATTTCTCGAGTCAAGTGA 1586
DB 1501 CCAAGAGATTAATGATTTCTCGAGTCAAGTGA 1533

RESULT 10
US-10-718-952-13
; Sequence 13, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACHARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 13
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-718-952-13

Query Match 83.3%; Score 1465.8; DB 19; Length 1533;
Beet Local Similarity 97.3%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 54 ATGTCATCGAGAAATTTAAGTTGAGTGTCTTAATGGAAGTACCCGAGACTGAGATT 113
DB 1 ATGTCATCGAGAAATTTAAGTTGAGTGTCTTAATGGAAGTACCCGAGACTGAGATT 60
QY 114 CAGTCGCTGACACTTACGAAACCAACCGAATTTGTTCAAGAGACAGAAATGACACTTAT 173
DB 61 CAGTCGCTGACACTTACGAAACCAACCGAATTTGTTCAAGAGACAGAAATGACACTTAT 120
QY 174 CAGTCGCTGACAAACCAAAATCTGTCAAAATACGAATTTAAACCAACATCATGTTCT 233
DB 121 CAGTCGCTGACAAACCAAAATCTGTCAAAATACGAATTTAAACCAACATCATGTTCT 180
QY 234 AAATTAGGGGTAATGCTGTGTGGGTGGAACCAACCGCTCAACCTCACCGGTGT 293
DB 181 AAATTAGGGGTAATGCTGTGTGGGTGGAACCAACCGCTCAACCTCACCGGTGT 240
QY 294 GTTATTTGCTAACCGAGAGGCTTTCATGGGCTACAAAGAGACAGATTTCAACAGCCAT 353
DB 241 GTTATTTGCTAACAGAGAGGCTTTCATGGGCTACAAAGAGACAGATTTCAACAGCCAT 300
QY 354 TACCTTGGCTCCCTCACCAAGCTCAGCTATCCGATGAGTGTGGTCTTCCAGGAGAGGA 413
DB 301 TACCTTGGCTCCCTCACCAAGCTCAGCTATCCGATGAGTGTGGTCTTCCAGGAGAGGA 360
QY 414 ATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTAAACCTGACAGATTTGTTGGG 473
DB 361 ATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTAAACCTGACAGATTTGTTGGG 420
QY 474 GGATGGATATCAGCAATGAACTGGTGTATGCTGATGCTGAGGCAAGGCTGTTGAC 533
DB 421 GGATGGATATCAGCAATGAACTGGTGTATGCTGATGCTGAGGCAAGGCTGTTGAC 480
QY 534 ATCGATTTGCAAGACAGTGTGAGGCTTACATGGAATCATGCTTCCACTCCCGGAATC 593
DB 481 ATCGATTTGCAAGACAGTGTGAGGCTTACATGGAATCATGCTTCCACTCCCGGAATC 540
QY 594 TATGACCCGGATTTGATTTGCTGCAACCAAGAGAGCGTGTCCCAACAGCTCATCAAGGCG 653
DB 541 TATGACCCGGATTTGATTTGCTGCAACCAAGAGAGCGTGTCCCAACAGCTCATCAAGGCG 600
QY 654 ACAAGCAAGACAGATTTCAAAATCATCAAAAGCATCAAGGCTTTAAGAGACCCACC 713
DB 601 ACAAGCAAGACAGATTTCAAAATCATCAAAAGCATCAAGGCTTTAAGAGACCCACC 660
QY 714 AAAGTGAACAAGTGGTGTGACTGTGAGCTGGCCAAACAGAGAGGTAGTAATTTGGT 773
DB 661 AAAGTGAACAAGTGGTGTGACTGTGAGCTGGCCAAACAGAGAGGTAGTAATTTGGT 720
QY 774 GTGGGCTTAATGACACCAATGAGAAATCTTTGGTCTGTGTGACAGAAATGAGGCTGAG 833
DB 721 GTAGGCTTAATGACACCAATGAGAAATCTTTGGTCTGTGTGACAGAAATGAGGCTGAG 780
QY 834 ATTTCTCTTCCACCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 893

```
Db      781 ATTTCTCTTCCACCTGTGATGCCATTCGTGTGATGGAATAATGTTCCATTATAT 840
Qy      894 GGAAGCCCTCAGAACTTTTGTACCAAGGCTGATGATTTTCCATCCGAGAACT 953
Db      841 GGAAGCCCTCAGAACTTTTGTACCAAGGCTGATGATTTTCCATCCGAGAACT 900
Qy      954 TTGATGTGTGAGATGATCAAGAGTGTGACAGCAAAATGAATCTGTGTGTGAT 1013
Db      901 TTGATGTGTGAGATGATCAAGAGTGTGACAGCAAAATGAATCTGTGTGTGAT 960
Qy      1014 TTCTTGTGGGGCTGTATCAAGCAACATCTATAGTCAGTACCAACATCTGGAAAC 1073
Db      961 TTTCTTGTGGGGCTGTATCAAGCAACATCTATAGTACCAACATCTGGAAAC 1020
Qy      1074 AATGATGTATGATCTTTTGGCTCCAAACTTTCCGTTCCAGGAATCTCCAAAGC 1133
Db      1021 AATGATGTATGATCTCTCGCTCCAAACTTTCCGTTCCAGGAATCTCCAAAGC 1080
Qy      1134 AACGTTGTGATGATGTCAGCAAGCAATGCCATCTCTATAGCTGGTGAACATCCA 1193
Db      1081 AACGTTGTGATGATGTCAGCAAGCAATGCCATCTCTATAGCTGGTGAACATCCC 1140
Qy      1194 GACCATGTTGTTTATTAATGATGTGCTTACGTAGGGGACAGCAAGAGCCATGAT 1253
Db      1141 GACCATGTTGTTTATTAATGATGTGCTTACGTAGGGGATGCAAGAGCCATGAT 1200
Qy      1254 GAGTACACTTCAGAGATATTATCATGGGTGGAAGAGCAACATGTTTTCACACATGC 1313
Db      1201 GAGTACACTTCAGAGATATTATCATGGGTGGAAGAGCAACATGTTTTCACACATGT 1260
Qy      1314 GAGGATTCCTCTTACCTGCTCTATTTATCTTGAGCTGGCTCTCTGCTGAGCTCAGC 1373
Db      1261 GAGGATTCCTCTTACCTGCTCTATTTATCTTGAGCTGGCTCTCTGCTGAGCTCAGC 1320
Qy      1374 ACTAGATCGAGTTTAAAGCTGAAAATGAGGAAAATTCACATCTATCCACCGATGGCT 1433
Db      1321 ACTAGATCGAGTTTAAAGCTGAAAATGAGGAAAATTCACATCTATCCACCGATGGCT 1380
Qy      1434 ACCATCTCAGCTACCTCACAAGGCTCTCTGTTTCCACGGGTACACCAAGTGTGAAT 1493
Db      1381 ACCATCTCAGCTACCTCACAAGGCTCTCTGTTTCCACGGGTACACCAAGTGTGAAT 1440
Qy      1494 GATTTGTCAAAGAGGCTGCAATGCTGGAATAACATATAGGGCTGTGTGTGATGGCC 1553
Db      1441 GATTTGTCAAAGAGGCTGCAATGCTGGAATAACATATAGGGCTGTGTGTGATGGCC 1500
Qy      1554 CCAGAGATATACATGATCTCGAGTACAAGTGA 1586
Db      1501 CCAGAGATATACATGATCTCGAGTACAAGTGA 1533

RESULT 11
US-10-718-952-15
; Sequence 15, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 9, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
```

```
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-718-952-15
Query Match      83.3%; Score 1465.8; DB 19; Length 1533;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy      54 ATGTCATCGAGAAATTTAAGGTGATGTCCTATGTGAGTACACCGAGCTGAGATT 113
Db      1 ATGTCATCGAGAAATTTAAGGTGATGATGTCCTAATGTGAGTACACCGAGCTGAGATT 60
Qy      114 CAGTCCTGTACAACTACGAAACCAACCGAATCTGTTACAGAAACGAAATGGCACTAT 173
Db      61 CAGTCCTGTACAACTACGAAACCAACCGAATCTGTTACAGAAACGAAATGGCACTAT 120
Qy      174 CAGTGAATGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATCTGTTCT 233
Db      121 CAGTGAATGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATCTGTTCT 180
Qy      234 AAATTAGGGGTATATCTGTGGGTGGGGTGAACCAACGCTCAACCTCAACGGTGT 293
Db      181 AAATTAGGGGTATATCTGTGGGTGGGGTGAACCAACGCTCAACCTCAACGGTGT 240
Qy      294 GTTATTCCTAACCGAGGGCATTTATGAGGCTACAAAGGAAAGATTTCAAGCCAT 353
Db      241 GTTATTCCTAACCGAGAGGGCATTTATGAGGCTACAAAGGAAAGATTTCAAGCCAT 300
Qy      354 TACTTGGCTCCCTACCCCAACGCTCAGCTATCCGATGAGTGGTCTTCCAGGAGAGNA 413
Db      301 TACTTGGCTCCCTACCCCAACGCTCAGCTATCCGATGAGTGGTCTTCCAGGAGAGNA 360
Qy      414 ATCTATGCCCCATTAAGAGCTGCTTCCATAGTTAACTCTGACGACATGTGTTGGG 473
Db      361 ATCTATGCCCCATTAAGAGCTGCTTCCATAGTTAACTCTGACGACATGTGTTGGG 420
Qy      474 GGATGGATATACGAACATGAACCTGGCTGATGTCATGGCCAGGCAAAAGTGTGAC 533
Db      421 GGATGGATATACGAACATGAACCTGGCTGATGTCATGGCCAGGCAAAAGTGTGAC 480
Qy      534 ATCGATTTGCAAGAGCTGAGGCTTATCATGATATCCATGCTTCCACTCCCGGAATC 593
Db      481 ATCGATTTGCAAGAGCTGAGGCTTATCATGATATCCATGCTTCCACTCCCGGAATC 540
Qy      594 TATGACCCGGATTTGATCTGTCACAACCAAGAGAGCGTCCCAACACCTCATCAAGGC 653
Db      541 TATGACCCGGATTTGATCTGTCACAACCAAGAGAGCGTCCCAACACCTCATCAAGGC 600
Qy      654 ACAAAAGCAAGAGCAATTCATAAATCATCAAAACATCAAGGCTTTAAGAAACCAAC 713
Db      601 ACAAAAGCAAGAGCAATTCATAAATCATCAAAACATCAAGGCTTTAAGAAACCAAC 660
Qy      714 AAAGTGAACAAGGCTGTGATCTGTCAGCTGTCACAACAGAGGTACAGTAATTTGGT 773
Db      661 AAAGTGAACAAGGCTGTGATCTGTCAGCTGTCACAACAGAGGTATAGCAATTTGGT 720
Qy      774 GTGGGCTTTAATGACACCATGAGATCTTTGCTGTGTGACAGAAATGAGCTGAG 833
Db      721 GTAGGCTTTAATGACACCATGAGATCTTTGCTGTGTGACAGAAATGAGCTGAG 780
Qy      834 ATTTCTCTTCCACTGTGTATGCAATGCTGTGTGTAAGAAATGTTCTTTCAATAT 893
Db      781 ATTTCTCTTCCACTGTGTATGCAATGCTGTGTGTAAGAAATGTTCTTTCAATAT 840
Qy      894 GGAAGCCCTCAGAACTTTTGTACCAAGGCTGATGATTTTCCATCCGAGAACT 953
Db      841 GGAAGCCCTCAGAACTTTTGTACCAAGGCTGATGATTTTCCATCCGAGAACT 900
Qy      954 TTGATGTGTGAGATGATCAAGAGTGTGACAGCAAAATGAATCTGTGTGTGAT 1013
Db      901 TTGATGTGTGAGATGATCAAGAGTGTGACAGCAAAATGAATCTGTGTGTGAT 960
Qy      1014 TTCTTGTGGGGCTGTATCAAGCAACATCTATAGTACCAACATCTGGAAAC 1073
```

|||||
Db 961 TTTCTTGGGGGCTGTATCAAGCCAACTCTATAGTTACACCATCTCGGAAAC 1020
Qy 1074 AATGATGATGATATCTTTCGGCTCCAACTTCCGTTCCAGAAATCTCCAAAGC 1133
Db 1021 AATGATGATGATATCTTTCGGCTCCAACTTCCGTTCCAGAAATCTCCAAAGC 1080
Qy 1134 AACGTGTGTATGATATGATGATCAAGCAATGCAATCTCTATAGCTGTGTGATCCA 1193
Db 1081 AACGTGTGTATGATATGATGATCAAGCAATGCAATCTCTATAGCTGTGTGATCCA 1140
Qy 1194 GACCAATGTTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
Db 1141 GACCAATGTTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Qy 1254 GAGTCACTTCAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
Db 1201 GAGTCACTTCAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Qy 1314 GAGATTCCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1373
Db 1261 GAGATTCCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Qy 1374 ACTAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1433
Db 1321 ACTAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Qy 1434 ACCATCTCAGCTACCTCAAGAGCTCTGATGATGATGATGATGATGATGATGATGAT 1493
Db 1381 ACCATCTCAGCTACCTCAAGAGCTCTGATGATGATGATGATGATGATGATGATGAT 1440
Qy 1494 GCATTGTCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553
Db 1441 GCATTGTCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Qy 1554 CCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586
Db 1501 CCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1533

RESULT 12

US-10-025-003-11
; Sequence 11, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Steile, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-11

Query Match 83.2%; Score 1464.2; DB 14; Length 1533;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 54 ATGTCATCGAGATTTTAAAGTGTCTTAATGTGAATGACACCGAGCTGAGATT 113
Db 1 ATGTCATCGAGATTTTAAAGTGTCTTAATGTGAATGACACCGAGCTGAGATT 60

Qy 114 CAGTCCGTGTAACAATAAGAAACCAACGAACTGTGTCAGAGAACAGGAATGGACCTAT 173
Db 61 CAGTCCGTGTAACAATAAGAAACCAACGAACTGTGTCAGAGAACAGGAATGGACCTAT 120
Qy 174 CAGTGGATGTCACCAACCAATCTGTCAAAATAGCAATTTAAACCAACATCTGTTCT 233
Db 121 CAGTGGATGTCACCAACCAATCTGTCAAAATAGCAATTTAAACCAACATCTGTTCT 180
Qy 234 AAATAGGGGATATGCTGTGGGTGGGATGAAACCAACGAGCTCAACCTCAGCGGTGT 293
Db 181 AAATAGGGGATATGCTGTGGGTGGGATGAAACCAACGAGCTCAACCTCAGCGGTGT 240
Qy 294 GTTATGCTAACCGAGAGGGATTTATGAGCTCAAAAGACAAAGATTCAA CAAGCCAT 353
Db 241 GTTATGCTAACCGAGAGGATTTATGAGCTCAAAAGACAAAGATTCAA CAAGCCAT 300
Qy 354 TACTTGGCTCCCTCACCAGGCTCAGCTATCCGAGTTGGGTCTTCCAGAGAGAGAA 413
Db 301 TACTTGGCTCCCTCACCAGGCTCAGCTATCCGAGTTGGGTCTTCCAGAGAGAGAA 360
Qy 414 ATCTATGCCCCATTCAGAGAGCTGCTCAATGTTAACCTGACGACATTTGTTGGG 473
Db 361 ATCTATGCCCCATTCAGAGAGCTGCTCAATGTTAACCTGACGACATTTGTTGGG 420
Qy 474 GGATGGATATCAAGCAATGAACCTGGCTGATGATGATGATGATGATGATGATGATGAT 533
Db 421 GGATGGATATCAAGCAATGAACCTGGCTGATGATGATGATGATGATGATGATGATGAT 480
Qy 534 ATCTATGTCAGAAAGCAATGAGGCTTACATGAAATCAATGCTTCCATCCCGGAATC 593
Db 481 ATCTATGTCAGAAAGCAATGAGGCTTACATGAAATCAATGCTTCCATCCCGGAATC 540
Qy 594 TATGACCCGATTCATGCTGTCACCAACCAAGAGAGCTGCAACAGTCAATCAAGAGG 653
Db 541 TATGACCCGATTCATGCTGTCACCAACCAAGAGAGCTGCAACAGTCAATCAAGAGG 600
Qy 654 ACAAGCAAGCAAGGTTCAACCAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 713
Db 601 ACAAGCAAGCAAGGTTCAACCAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 660
Qy 714 AAAGTGAACAAGGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
Db 661 AAAGTGAACAAGGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy 774 GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGTGTCAGAGAAATGAGGCTGAG 833
Db 721 GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGTGTCAGAGAAATGAGGCTGAG 780
Qy 834 ATTTCTCTTCAACCTTGTATGCAATGCTTGTGTATGAAATGTTCTTTCTATAT 893
Db 781 ATTTCTCTTCAACCTTGTATGCAATGCTTGTGTATGAAATGTTCTTTCTATAT 840
Qy 894 GGAAGCCTCAGACACCTTTGTACAGAGGCTGATGATGATGATGATGATGATGATGATGAT 953
Db 841 GGAAGCCTCAGACACCTTTGTACAGAGGCTGATGATGATGATGATGATGATGATGATGAT 900
Qy 954 TTGATGTTGAGATGATCTTCAAGAGTGTGAGACCAAAATGAATCTGTGTGTTGAT 1013
Db 901 TTGATGTTGAGATGATCTTCAAGAGTGTGAGACCAAAATGAATCTGTGTGTTGAT 960
Qy 1014 TTCTTGTGGGGGCTGTATCAAGCCAACTATAGTCAAGTTACCAACATCTGGAAC 1073
Db 961 TTCTTGTGGGGGCTGTATCAAGCCAACTATAGTCAAGTTACCAACATCTGGAAC 1020
Qy 1074 AATGATGATGATCTTTCGGCTCCAACTTCCGTTCCAGAAATCTCCAAAGC 1133
Db 1021 AATGATGATGATCTTTCGGCTCCAACTTCCGTTCCAGAAATCTCCAAAGC 1080
Qy 1134 AACGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
Db 1081 AACGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140

OY	1194	AACCAATGTTGTTGTAATTAAGAAATGCGCTTAACGAGGGACACGAAGAAGCCATGGAT	1255
Db	1141	GACCATTGTTGTTGTTATTAAGATATGTGCTCTTAAGAGGAGATAGGAAGAAGGCCATGGAT	1200
OY	1254	GAGTACACTTCAGAGATATATCATGGGTGGAAAGACACCATTTGTTTTCACAACAACATGC	1313
Db	1201	GAGTACACTTCAGAGATATATCATGGGTGGAAAGACACCATTTGTTTTCACAACAACATGT	1266
OY	1314	GAGGATTCCTCTTAGCTGCTGCTCTTAATTATCTTGGACTTGGTCTCTTCTTGCTGAGCTCAGC	1373
Db	1261	GAGGATTCCTCTTATTAGCTGCTCTTAATTATCTTGGACTTGGTCTCTTCTTGCTGAGCTGAGC	1320
OY	1374	ACTGAATCGAGTTTAAAGCTGAAATATAGGGAATTCACATCAATTCACACCCAGTTGCT	1433
Db	1321	ACTGAATTCAGTTTAAAGCTGAAATATAGGGAATTCACATCAATTCACACCCAGTTGCT	1380
OY	1434	ACCAATCCTCAGCTACCTCACCAAGGCTCCTCTGGTTTCCACCGGGATACCAAGTGGTAAT	1493
Db	1381	ACCAATTCACAGTATCTGACCAAGGCTCCTCTGGTTTCCACCGGGATACCAAGTGGTAAT	1440
OY	1494	GCATTGTCAAAAGCAGCGTGCATATGCTGGAAAAATATATAGAGGGCTTGTGTGGATTGGCC	1553
Db	1441	GCATTGTCAAAAGCAGCGTGCATATGCTGGAAAAATATAGAGGGCTTGTGTGGATTGGCC	1500
OY	1554	CCAGAGAAATACATGATCTTCGAGATCAAGAAGA	1586
Db	1501	CCAGAGAAATACATGATCTTCGAGATCAAGAAGA	1533

```

RESULT 13
US-10-718-952-11
; Sequence 11, Application US/10718952
; Publication NO. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYABAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-718-952-11

```

Query Match	83.2%	Score 1464.2	DB 19	Length 1533
Best Local Similarity	97.2%	Pred. No. 0		
Matches 1490	Conservative	0	Mismatches 43	Indels 0
				Gaps 0

QY	54	ATGTTTCATCGAAATTTTAAGGTTGAGTGTCTTAATGGAAGTA	CACCGAGACTTGAGATT	113
Db	1	ATGTTTCATCGAAATTTTAAGTGAAGAGTCTTAATGGAAGTA	CACCGAGACTTGAGATT	60
QY	114	CAGTCCGTTGACAACTACGAAACCA	CCGAACTTGTTACGGAACAGGAATGGCACTTAT	173
Db	61	CAGTCCGTTGACAACTACGAAACCA	CCGAACTTGTTACGGAACAGGAATGGCACTTAT	120
QY	174	CAGTGAATGTTCAAAACCCAAATCTGTCAATACGAATTTAAACCA	CATCATGTTCTT	233
Db	121	CAGTGAATGTTCAAAACCCAAATCGGTCAACTACGAATTTAAACCA	CATCATGTTCTT	180
QY	234	AAATTTAGGGGTAAATGCTTTGGGTTGGGGTGGAAACCA	CGGCTCAACCTTACCGGTGGT	293

Db	181	AAATTGGGGGGTGATGCTTGTGGGTTGGGGGTGAAAACAAAGCGCTCAACCTTCAACCGGTGT	240
OY	294	GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGGACAAGATTCAACAAGCCAA	353
Db	241	GTTATTGCTTAACAGAGGAGACATTTTCATGGGCTACAAAGGACAAGATTCAACAAGCCAA	300
OY	354	TACTTTGGCTCCCTCAACCAAGCCTCAGCTATCCGAGTTGGGTCCTTCAAGGAGAGAA	413
Db	301	TACTTTGGCTCCCTCAACCAAGCCTCAGCTATTCGAGTTGGATCTTTCAGAGGAGAGAA	360
OY	414	ATCTATGCCCATTCAGAGGCTGCTTCCATATGGTTAAACCTGACGACATGTGTTTGGG	473
Db	361	ATCTATGCCCATTCAGAGGCTGCTTCCATATGGTTAACTCTGACGACATGTGTTTGGG	420
OY	474	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGCGCAGAGGGCAAGAGTGTTTGAC	533
Db	421	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGCGCAGAGGGCAAGAGTGTTTGAC	480
OY	534	ATCGATTTGCAAGAACAGTTGAGGCTTTACATGGAATCATTGCTTCCACTCCCGGAATC	593
Db	481	ATCGATTTGCAAGAACAGTTGAGGCTTTACATGGAATCATTGCTTCCACTCCCGGAATC	540
OY	594	TATGACCCGGGATTTTCATTTGCTGCCACCAAGAGAGGCTGCCAACAAGCTATCAAGGCG	653
Db	541	TATGACCCGGGATTTTCATTTGCTGCCACCAAGAGAGGCGTGCACAAAGTATTAAGGGC	600
OY	654	ACAAAGCAAGAGCAAGTTCAACAATATCATCAAGACATCAAGGGGTTTAAAGAAAGCCAC	713
Db	601	ACAAAGCAAGAGCAAGTTCAACAATATCATCAAGACATCAAGGGGTTTAAAGAAAGCCAC	660
OY	714	AAAGTGACAAAGGTGGTTGTACTGTGACCTGCCAAACAGAGAGGTACAGTAATTTGGTT	773
Db	661	AAAGTGACAAAGGTGGTTGCTCTGTGACCTGCCAAACAGAGAGGTATAGCAATTTGGTT	720
OY	774	GTGGGCTCTTAATGACACCATGAGAAATCTCTTGGCTCTGTGACACAGAAATGAGAGCTGAG	833
Db	721	GTAGGCTCTTAATGACACCATGAGAAATCTCTTGGCTCTGTGACACAGAAATGAGAGCTGAG	780
OY	834	ATTTCTCCCTTCCACTTGTATGCATTTGCTGTGTATGAGAAATGTTCCTTTCAATTAAT	893
Db	781	ATTTCTCCCTTCCACTTGTATGCATTTGCTGTGTATGAGAAATGTTCCTTTCAATTAAT	840
OY	894	GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTGATCTTTGCCATCGCAGAGAAC	953
Db	841	GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTGATCTTTGCCATCGCAGAGAAC	900
OY	954	TTGATTTGGTGAGATGACTTCACAGAGTGTCAACCAATGAAATCTGTGTGGTTGAT	1013
Db	901	TTGATTTGGTGAGATGACTTCACAGAGTGTCAACCAATGAAATCTGTGTGGTTGAT	960
OY	1014	TTCTGTGGGGGGCTGGATTCAAAGCCAAATCTATAGCTAGTTTCAACCAATCTGGGAAC	1073
Db	961	TTCTGTGGGGGGCTGGATTCAAAGCCAAATCTATAGTATGTTTCAACCAATCTGGGAAC	1020
OY	1074	AATGATGGTATGAATCTTTTGGGCTCCAAACCTTTCGGTTCCAGAGAAATCTCCAAAGC	1133
Db	1021	AATGATGGTATGAATCTCTGGGCTCCAAACCTTTCGGCTCCAGAGAAATCTCCAAAGC	1080
OY	1134	AAAGTTGGTGAATGATGATGATCAGCAAGATGCAATCCTATAGAGCTGTGAAACATCCA	1193
Db	1081	AAAGTTGGTGAATGATGATGATCAGCAAGATGCAATCCTATAGAGCTGTGAAACATCCC	1140
OY	1194	GACCATGTGTTGTTAATTAAGTATGAGCTTAAGCGGAGAGGAGAGGCAATGAT	1253
Db	1141	GACCATGTGTTGTTAATTAAGTATGAGCTTAAGCGGAGAGGAGAGGCAATGAT	1200
OY	1254	GAGTACACTTCAGAGATTAATTCATGGGTGAAAGAGCAATGTTGTTTGCAACAACATGCG	1313
Db	1201	GAGTACACTTCAGAGATTAATTCATGGGTGAAAGAGCAATGTTGTTTGCAACAACATGCT	1260
OY	1314	GAGGATTCCTCTTAGCTGCTCTTAATCTTGGACTTGGCTCTTCTTGCTGAGCTGAG	1373
Db	1261	GAGGATTCCTCTTAGCTGCTCTTAATCTTGGACTTGGCTCTTCTTGCTGAGCTGAG	1320

QY 1374 ACTAGATCGAGTTTAAAGCTGAAATGAGGAAATTCCTCATTCACCCAGTTGCT 1433
DB 1321 ACTAGATCGAGTTTAAAGCTGAAATGAGGAAATTCCTCATTCACCCAGTTGCT 1380
QY 1434 ACCATCTCAGCTACCTCAGCAAGGCTCTCTGTTCCACCGGGTACACCACTGTGAAT 1493
DB 1381 ACCATCTCAGCTACCTCAGCAAGGCTCTCTGTTCCACCGGGTACACCACTGTGAAT 1440
QY 1494 GATTTGTCAAAGCAGCTGCAATGCTGGAACCATTAAGAGGGCTGTGTTGGATTGGCC 1553
DB 1441 GATTTGTCAAAGCAGCTGCAATGCTGGAACCATTAAGAGGGCTGTGTTGGATTGGCC 1500
QY 1554 CCAGAGAAATTAACATGATTTCTCGAGTACAACTGA 1586
DB 1501 CCAGAGAAATTAACATGATTTCTCGAGTACAACTGA 1533

RESULT 14

US-10-424-599-12022
; Sequence 12022, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21 (53223) B
; CURRENT APPLICATION NUMBER: US/10-424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12022
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1
US-10-424-599-12022

Query Match 73.7%; Score 1296.4; DB 18; Length 2018;

Best Local Similarity 88.6%; Pred. No. 0;

Matches 1405; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 47 AAAAAATATGTCATCGAGAAATTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGAC 106
DB 91 AAAAAAATGTCATCGAGAGTTTCAAGGTTGAGTGTCTTAATGTAAGTACCGAGAC 150
QY 107 TGAGATTCAGTCCGTTGACAACTACGAAACCAACCTTTGTCAGAGAACAGAAATG 166
DB 151 TGAGATTCAGTCCGTTGACAACTACGAAACCAACCTTTGTCAGAGAACAAATATG 210
QY 167 CACTATCAGTGGATTTGCAAAACCAATCTGTCAAAATGCAATTTAAAAACAATCCA 226
DB 211 CACTATCAGTGGATTTGCAAAACCAATCTGTCAAAATGCAATTTAAAAACAATCCA 270
QY 227 TCTTCTTAATTAAGGGGTAACTGTTGGGTGGGTGGAACAACGGCTCAACCTCAC 286
DB 271 TCTTCTTAATTAAGGGGTAACTGTTGGGTGGGTGGAACAACGGCTCAACCTCAC 330
QY 287 CGGTGGTGTATTTGCTAACCGAGAGGGCAATTCATGGGCTACAAAGAACAAATTCACA 346
DB 331 TGGTGGTGTATTTGCTAACCGAGAGGGTATTTCAATGGGCAACAAAGAACAAATTCACA 390
QY 347 AGCCAAATTAATTTGGCTCCCTCAACCCAGCTCAGCTATCCAGTTGGGTCTTCAGGG 406
DB 391 AGCCAAATTAATTTGGGTCCCTCAACCCAGCTCAGCTATCCAGTTGGGTCTTCAGGG 450
QY 407 AGAGGAATCTATGCCCCCTTCAAGAGCCTGCTTCCAAATGTTAAACCTGAGACATTTGT 466
DB 451 GGAAGAAATATATGCTCAATTCAGAGCCTGCTTCCAAATGTTAAACCTGAGATGTTGT 510

QY 467 GTTTGGGGGATGGAGATATACGAAACATGAACTGTGATGCCATATGCGCAGGGCAAGGT 526
DB 511 GTTTGGGGGATGGAGATATACGAAACATGAACTGTGATGCCATATGCGCAGGGCAAGGT 570
QY 527 GTTTGACATGATTTTGCAGAAAGCAGTTGAGGCCCTTAATGGAATCCATGCTTCCATCCC 586
DB 571 GTTTGATATGCACTCTGCAAGAAACAGTTGAGGCCCTTAATGGAATCCATGCTTCCATCCC 630
QY 587 CGGAATCTATGACCCGGATTTTCATTTGCTGCCAACCAAGAGGAGCGTGCCAAAGATCAT 646
DB 631 CGGAATCTATGACCCGGATTTTCATTTGCTGCCAACCAAGAGGAGCGTGCCATTAAGTAT 690
QY 647 CAAGGGCAAAAGCAAGCAAGCAAGTTCAACAAATCATCAAGACATCAAGCGTTTAAGA 706
DB 691 CAAGGGCAAAAGCAAGCAAGCAAGTTCAACAAATCATCAAGACATTAAGAGATTCAAGA 750
QY 707 AGCCACCAATGGAACAAAGGTGTTGTAATCTGGAATGCGCAACACAGAGAGTACAGTAA 766
DB 751 AAGCACTAAGGTTGCAAGAGGTGTTGTAATCTGGAACAGCCCAACACAGAGAGTACAGCA 810
QY 767 TTTGGTTGGGACCTTAATGACACCATGAGAAATCTTTGCTGCTGTCGACAGAAATGA 826
DB 811 GGTAGTTGTGGAGACTTAACGACACCATGAGAAACCTCTTGCTTCTTGACAGAGAAATGA 870
QY 827 GGTGAGATTTCTCCTTCACCTTGATGATGCAATGCTTGTGTTATGAAAAATGTTCTTT 886
DB 871 GGTGAGATTTCTCCTTCACCTTGATGATGCAATGCTTGTGTTATGAAAAATGTTCTTT 930
QY 887 CATTAATGGAAGCCTCGAACACTTTTGTACAGGGGTGATTAATCTTGCAATCGCAG 946
DB 931 CATCAATGGAAGCCTCGAACACTTTTGTCCAGGAGCTCAATGATCTGCGCATCAAAAG 990
QY 947 GAACACTTGTATGGTGGAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTT 1006
DB 991 GAATAGTTGATTTGAGAGAGATGACTTAAAGTGTGACAGCAAAATGAATGAGTGT 1050
QY 1007 GGTGATTTCTTGTGGGGGCTGTGATCAAGCCAACTATATGCTCAATCAACATCT 1066
DB 1051 GGTGATTTCTTGTGAGAGAGATGACTTAAAGTGTGACAGCAAAATGAATGAGTGT 1110
QY 1067 GGGAAACAAATGATGATTAATTTTGGCTCCACAACTTCCGTTCCAGAGAAATCTC 1126
DB 1111 GGGAAACAAATGATGATTAATTTTGGCTCCACAACTTCCGCTCGAAGAAATTTTC 1170
QY 1127 CAAGAGCAACGTTGTGATGATATGATCAAGAGCAATCCATCTCTATGAGCCTGTGTA 1186
DB 1171 CAAGAGCAATGTTGTGACGATGATGATCAAGAGCAATCCATCTCTATGAGCCTGTGTA 1230
QY 1187 ACATCCAGACCATGTTGTTGTTATTAAGTATGTGCTTACGTTAGGGGACAGCAAGAGAC 1246
DB 1231 ACATCCAGACCATGTTGTTGTTATTAAGTATGTGCTTACGTTAGGGGACAGCAAGAGAC 1290
QY 1247 CATGATGATGATCACTTCAAGATATTTCAATGGGTGGAAGACCAATGTTTTCACAA 1306
DB 1291 CATGATGATGATCACTTCTTCAAGATATTTCAATGGGTGGAAGACCAATGTTGTTCAAA 1350
QY 1307 CACATGAGAGATTTCCCTCTTATGCTGCTCTATTAATCTTGAACCTGTGCTTCTTCTGTA 1366
DB 1351 CACATGATGATGATCACTGTTAGCTGCTCTATTAATCTTGAACCTGTGCTTCTTCTGTA 1410
QY 1367 GCTCAGCACTAGATGAGATTTAAAGCTGAAATGAGGAAATTTCCATCAATTCACCC 1426
DB 1411 GCTTACGACTGGAATCGATTTAAAGCTGAAATGAGGAAATTTCCATCAATTCACCC 1470
QY 1427 AGTTGCTACATCTCAGCTACCTACCAAGGCTCTTGGTTCCACGGGTACACAGT 1486
DB 1471 AGTTGCTACATCTCAGCTACCTACCAAGGCTCTTGGTTCCACGGGTACACAGT 1530
QY 1487 GGTGATGATTTGCAAGAGCGGTGCAATGCTGGAACCAATATGAGGGCTTGTGTTGG 1546
DB 1531 GGTATATCACTGTCAAGAGCGGTGCAATGCTGGAACCAATATGAGGGCTTGTGTTGG 1590
QY 1547 ATTGCCCCCAGAGATTAACATGATTTCTGAGTACAAAGTGAACATGGGACCGAAGAAATTA 1606

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 21:02:47 ; Search time 300.905 Seconds
(without alignments)
9570.620 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctctatctcctcttctgt.....aatgcagataattctgtc 1760

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfileseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	994.2	56.5	1959	4	US-09-727-628-1
2	984	55.9	1931	3	US-09-118-442-10
3	984	55.9	1931	3	US-09-677-064-10
4	384.8	21.9	1578	4	US-09-248-796A-3131
5	349.4	19.9	1231	3	US-09-397-787-34
6	340.2	19.3	1602	4	US-09-734-237B-72
7	340	19.3	1605	4	US-09-734-237B-74
8	160.6	9.1	3546	3	US-09-118-442-15
9	160.6	9.1	3546	3	US-09-677-064-15
10	159	9.0	3546	3	US-09-118-442-14
11	159	9.0	3546	3	US-09-677-064-14
12	89.4	5.1	77626	4	US-09-949-016-12608
13	62.2	3.5	294	4	US-09-313-294A-4684
14	48.4	2.8	1335	4	US-09-902-540-5256
15	48.4	2.8	34316	4	US-09-902-540-1257
16	44.6	2.5	1141	4	US-09-806-708B-22
17	37.6	2.1	7218	1	US-08-232-463-14
18	37.6	2.1	601	4	US-09-949-016-59935
19	37.2	2.1	601	4	US-09-949-016-117235
20	37.2	2.1	601	4	US-09-949-016-117304
21	37.2	2.1	77388	4	US-09-949-016-13496
22	37.2	2.1	115388	4	US-09-949-016-14881
23	37.2	2.1	127771	4	US-09-949-016-14882
24	36.6	2.1	512	4	US-09-270-767-5181
25	36.6	2.1	512	4	US-09-270-767-20463
26	36.2	2.1	1098	3	US-09-221-017B-928
27	35.6	2.0	251769	4	US-09-949-016-13185

C	28	35.6	2.0	251769	4	US-09-949-016-13186	Sequence 13186, A
C	29	35.6	2.0	266748	4	US-09-949-016-13187	Sequence 13187, A
C	30	35.6	2.0	266748	4	US-09-949-016-13188	Sequence 13188, A
C	31	35.2	2.0	1584	4	US-09-602-777A-419	Sequence 419, App
C	32	35.2	2.0	2310	5	PCT-US95-0724A-4	Sequence 4, Appl1
C	33	35.2	2.0	2310	5	PCT-US95-0724A-4	Sequence 4, Appl1
C	34	35.2	2.0	3387	5	PCT-US95-0724A-5	Sequence 5, Appl1
C	35	35.2	2.0	3387	5	PCT-US95-0724A-5	Sequence 5, Appl1
C	36	35.2	2.0	5000	3	US-09-104-070-1	Sequence 1, Appl1
C	37	35	2.0	505	4	US-09-621-976-15639	Sequence 15639, A
C	38	35	2.0	2025	4	US-09-107-532A-2547	Sequence 2547, Ap
C	39	34.8	2.0	399	4	US-09-621-976-8976	Sequence 8976, Ap
C	40	34.8	2.0	601	4	US-09-949-016-14892	Sequence 14892, A
C	41	34.8	2.0	2518	3	US-09-433-699-3	Sequence 3, Appl1
C	42	34.8	2.0	2527	4	US-09-949-016-4169	Sequence 4169, Ap
C	43	34.8	2.0	13489	4	US-09-949-016-15911	Sequence 15911, A
C	44	34.6	2.0	291	4	US-09-313-294A-518	Sequence 518, App
C	45	34.4	2.0	1446	4	US-09-902-540-5772	Sequence 5772, Ap

ALIGNMENTS

RESULT 1									
US-09-727-628-1									
Sequence 1, Application US/09727628									
Patent No. 6791013									
GENERAL INFORMATION:									
APPLICANT: Armstrong, Katherine									
APPLICANT: Hey, Timothy D									
APPLICANT: Folkerts, Otto									
APPLICANT: Smith, Kelley A									
APPLICANT: Hopkins, Nicole L									
TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER									
FILE REFERENCE: 50597									
CURRENT APPLICATION NUMBER: US/09/727, 628									
PRIOR FILING DATE: 2000-12-01									
PRIOR APPLICATION NUMBER: US 60/168, 612									
PRIOR FILING DATE: 1999-12-02									
NUMBER OF SEQ ID NOS: 3									
SOFTWARE: PatentIn Ver. 2.0									
SEQ ID NO 1									
LENGTH: 1959									
TYPE: DNA									
ORGANISM: Zea mays									
FEATURE:									
NAME/KEY: CDS									
LOCATION: (137)..(1699)									
US-09-727-628-1									
Query Match									
Best Local Similarity 77.8%; Score 994.2; DB 4; Length 1959;									
Matches 1200; Conservative 0; Mismatches 343; Indels 0; Gaps 0;									
QY	54	AGTTTCATGAGAAATTTTAAAGTTGAGTCTCTATGTGAATGACACCGAGCTAGATT	113						
DB	137	AGTTTCATGAGAAATTTTAAAGTTGAGTCTCTATGTGAATGACACCGAGCTAGATT	196						
QY	114	CAGTCCGTTGACAACTTGAACCAACCGAATCTTGTTCACGAAACGAAATGACACTAT	173						
DB	197	GAGTCCGTTGACAACTTGAACCAACCGAATCTTGTTCACGAAACGAAATGACACTAT	256						
QY	174	CAGTCCGTTGACAACTTGAACCAACCGAATCTTGTTCACGAAACGAAATGACACTAT	233						
DB	257	CAGTCCGTTGACAACTTGAACCAACCGAATCTTGTTCACGAAACGAAATGACACTAT	316						
QY	234	AAATTAGGAGTAAATCTTGTGTTGAGTGAACCAACCGCTCAACCTGACCGTGT	293						
DB	317	AAATTAGGAGTAAATCTTGTGTTGAGTGAACCAACCGCTCAACCTGACCGTGT	376						
QY	294	GTTATTTGCTAACGAGAGGAGGATTTTCATGAGGCTTACAAAGCAAGTTCACAAAGCAAT	353						
DB	377	GTTATTTGCTAACGAGAGGAGGATTTTCATGAGGCTTACAAAGCAAGTTCACAAAGCAAT	436						

```
OY 354 TACTTTGGCTCCCTACCCCAAGCTCATCCAGTTGGGTCCTTCCAGGGAAGAGAA 413
DB 437 TACTAGGCTCCCTCAACCCAGGCTTCCACATCAGATGAGCACTCAACAGGGAAGAG 496
OY 414 ATCTATGCCCATTCAGAGAGCTGCTTCCAAATGGTTAACTTGAAGCAATGGTTGGG 473
DB 497 ATATATGCGCGCTTCAAGAGCTCTCTAACCATGGTGAACCCAGAGCAACTGTGTGGGA 556
OY 474 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGAGGCAAAAGTTTGAAC 533
DB 557 GGCTGGGACATCAGCAGCAATGAACCTGGAGATGCCATGACAGAGGCGCAAGGTGTGAGAC 616
OY 534 ATCGATTTGCGAAGAGCTTGAAGGCTTACATGGAATCCATGCTTCCATCCCGGAATC 593
DB 617 ATTTGACCTCGAAGAGAGCTCAGGCGCTTACATGAGATCCATGGTGCACCTTCCCGGTGC 676
OY 594 TATGACCCCGAATTTCAATGCTGCCAACAAGAGAGCGTGCACAACGCTCATCAAGGCG 653
DB 677 TATGATCCGGACTTCAATGCGCGCTAACAGGGGCTCTGTCGCAACAAATGTCATCAAGGCG 736
OY 654 ACCAAGCAAGAGCAAGTTCAACAAATCATCAAAAGACATCAAGCGCTTTAAGGAAGCCACC 713
DB 737 ACCAAGAAAAGAACAGGTGAGAGACATCAAGATATCAGGAGGTTTAAAGAGAAGAAC 796
OY 714 AAAAGTGAACAAGGTGTTGTAAGTGAAGTGCACAAGAGAGGTGACAGTATTTGGTT 773
DB 797 AAAAGTGAACAAGGTGTTGTAAGTGAAGTGCACAAGAGGTGACAGTATTTGGTT 856
OY 774 GTGGGCTTAAATGACCAATGAGAAATCTCTTGGCTGTGTGAGACAGAAATAGAGCTGAG 833
DB 857 GCTGGCTCAACGACACATGAGAAATCTGTGGCTGTGTGAGACAGAAATAGAGCTGAG 916
OY 834 ATTTCTCTTCAACCTGTAATGCCATTTGCTGTGTGTAAGAAATGTTCTTCAATTAAT 893
DB 917 ATCTGCGCATCAACATATATGCAATTTGCTGTGTGTAAGAAATGTTCTTCAATTAAT 976
OY 894 GGAAGGCTTCAGAACTTTTGTACAGAGGCTGATGATCTTGTCCATCGCAGAAACACT 953
DB 977 GGAAGGCTTCAGAACTTTTGTACAGAGGCTGATGATCTTGTCCATCGCAGAAACACT 1036
OY 954 TTGATTTGTGAGAGATGACTTCAAGAGTGTCAAGCCAAATGAATCTGTGTGTTGAT 1013
DB 1037 CTGATCGGTGTGAGACACTTCAAGAGTGTGAGCCAAAGATGAATCGGTCTGTGTTGAT 1096
OY 1014 TTCTTTGTGGGGCTGTATCAAGCCAAATCTATATGTCAAGTTCAACCATCTGGGAAC 1073
DB 1097 TTTCTTTGTGTGTGAGAAATGAAGCCACCTGATTTGAGCTACCAACCACTTGGGAAC 1156
OY 1074 AATGATGTATGATCTTTCGAGCTCCACAAACTTTGCTCCAAAGAAATCTCAAGAGC 1133
DB 1157 AACGAGCGGATGAACCTGTCTGCTCCAAACATTAAGTCCAAAGAGATCTCCAAAGAGC 1216
OY 1134 AACGTTTGTATGATATGATGTCAACAGCAATGCCATCTCTATGAGCGTGTGAACATCCA 1193
DB 1217 AACGTTTGTATGATATGATGTCAACAGCAATGCCATCTCTATGAGCGCGGAGAGATCCC 1276
OY 1194 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGACCATGTGAT 1253
DB 1277 GATCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGACCATGTGAT 1336
OY 1254 GATGACCTTCAAGATATTCATGGGTGGAAGAGAGACCATGTTTGTGCAACACATGAC 1313
DB 1337 GATGACACATCAAGATATTCATGGGTGGAAGAGAGACCATGTTTGTGCAACACATGAC 1396
OY 1314 GAGATTTCCCTCTTACGCTCTCTATTAATCTTGAATGATCCCTTCTTGTGAGCTCAAC 1373
DB 1397 GAGGATCGCTCTCTGCGGACCGGATCATCTGATCTGTGCTCTGTGCTGAGCTCAAC 1456
OY 1374 ACTAGATGAGTTTAAAGTGAAGAAATGAGGAAATTTCACTCATTTCCACCAAGTTGCT 1433
DB 1457 ACCAGGATCAAGTTAAACCTGAGGGAAGGAGCAAGTTCCACTTTCACACCGGTGAGCC 1516
```

```
OY 1434 ACCATCTCAGCTACCTCAGCAAGGCTCTGTGTTCCACCGGGTACCAAGTGTGAT 1493
DB 1517 AACATCTTAGGTACCTCAGCAAGGACCACTGGTTCCACCGGACACCGGTGGTGAAC 1576
OY 1494 GCATTGTCAAAGCAGCGTGCATATGTGGAAAACATTAATGAGGGCTGTGTGATTTGCC 1553
DB 1577 GCTCTTGCAAAAGCAGAGGCGCATGTGGAGAACATCAAGAGAGGCTGTGGTGGCTGGCC 1636
OY 1554 CCAGAGATTAATGATATCTCGATGATCAAGTGAAGCATGGGAC 1596
DB 1637 CCAGAGATTAATGATATCTCGATGATCAAGTGAAGCATGGGAC 1679

RESULT 2
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phyrate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10

Query Match 55.9%; Score 984; DB 3; Length 1931;
Best Local Similarity 76.1%; Pred. No. 3.5e-296;
Matches 1212; Conservative 0; Mismatches 380; Indels 0; Gaps 0;

OY 5 TCTTTATTCCTTTTGTATTTTCATTCATTTTAAATCTTTTGAAGAAATTAATGTTATGCA 64
DB 50 TCGTACCTCGCTTGCATTCATTCGAAAAGAGGAGAAAGCAAGATGTTTCATGCA 109
OY 65 GAATTTTAAGGTTGATGATCTTAATGTGAAGTACACGAGACTGAGATTCAATCCGTGTA 124
DB 110 GAGCTTCCGCTCGAGAGCGCCCACTGCTGAGGCGCCGATGAGATCGAGTCCGAGTA 169
OY 125 CAATTCAGAAACCAACGAACTTGTTCACAGAAACAGAAATGCACTATCAATGATGATGT 184
DB 170 CCGTACGACAGACAGAGAGCTGTACACAGAGGCAAGAGCGGCTCACGCTGAGTGT 229
OY 185 CAATCCCAATATGTCAATATGCAATTTAAACCAACATTCATGTTCTTAAATTAATGGGGT 244
DB 230 CCGCCCAAGTCCGTCAGATCACTTCGGAACCGAACCAGCCGCTCCCAAGCTGGGGGT 289
OY 245 AATGCTGTGGGTGGGGTGAAGAAACAGAGGCTCAACCTCAACCGGTGGTATTTGTTAA 304
DB 290 GATGCTGTGGGTGGGGTGAAGAAACAGAGGCTCAACCTCAACCGGTGGGTATTTGTTAA 349
OY 305 CCGAGAGGAGCTTTCATGAGGCTTCAAGAGCAAGATTCACAGCAATTAATTTGAGCTC 364
DB 350 CAGGAGAGGAGATCTATGAGGCGACCAAGGATGACAGAGCACTACTAGGCTC 409
OY 365 CCTCACCAAGCTTCAAGCTATTCAGATTGGGCTCTTCCAGGAGAGGAAATCTATGCCCC 424
DB 410 CTTCAACCAAGCTTCAACATCAGATCGGCGAGCTTACCAACGAGGAGAGATCTATGCGCC 469
```

QY 425 ATTCAAGAGCGCTTCAATGGTTAACTGACGCAATTTGTTGGGGATGGATAT 484
 DB 470 GTTCAAGAGCGCTTCCATAGTGAACCGACGCAATTTGTTGGGGATGGATAT 529
 QY 485 CAGCAACATGAACCTGGCTGATGTCATGGCCAGGGGAAAGTGTGACATCGATTGCA 544
 DB 530 TAGCAACATGAACCTGGCCGATCTCATGACAGGGCCAGAGTGTGATTTGACCTGCA 589
 QY 545 GAAGCATGGAGCGCTTCAATGGATTCATGCTTCCATCTCCGGAATCTATGACCCTGA 604
 DB 590 GAAGCATGGAGCGCTTCAATGGATTCATGCTTCCATCTCCGGAATCTATGATCCGGA 649
 QY 605 TTTCAATTTGCTGCAACCAAGAGAGGCTCCAAACGATCATTAAGGCAACAAAGCAAA 664
 DB 650 CTTCATCGGGCTTACCAAGAGGCTCTCGCGCAACAGTGTATCAAGGGCAACAAAGAA 709
 QY 665 GCAAGTTCAACAATCATCAAGACATCAAGGCGTTTAAAGAAACCAACCAAGTGAACA 724
 DB 710 ACAAGTGAAGCATCATCAAGATTCAGAGGATTTAAAGAGAAACAAAGTGAACA 769
 QY 725 GGTGTTTACTGTGAGCTGCCAACAAGAGATCAAGTAAATTTGGTGTGGCCTTAA 784
 DB 770 GATAGTTGTGTGAGCTGCAACAACGAAAGTATAGCAATGTGTGGCTGTCTCAA 829
 QY 785 TGACACCATGAGAAATCTTGGCTGTGTGACACAAATGAGGCTGAGATTTCTCTTC 844
 DB 830 CCAACGATGAGAAATCTTGGCTGTGTGACACAAATGAGGCTGAGATTTCTCTTC 889
 QY 845 CACCTTGTATGCAATGCTTGTGTATGGAATAATGTTCTTCAATTAATGAAGCCCTCA 904
 DB 890 AACCATATGCTATGCTGTGTATGGAATAATGTTCTTCAATTAATGAAGCCCTCA 949
 QY 905 GAACCTTTTGTACAGAGGCTGATGATCTTGGCATCGGAGAACATTTGATTTG 964
 DB 950 GAACCTTTTGTACAGAGGCTGATGATCTTGGCATTAATAAACAATGCTTATGTTG 1009
 QY 965 AATATGCTTCAAGAGGCTGACACCAAAATGAAATGTTGTTGTTGATTTCTTGTGG 1024
 DB 1010 TACAGCTTCAAGAGGCTGACACCAAAATGAAATGTTGTTGTTGATTTCTTGTGG 1069
 QY 1025 GAGCTGTATCAAGGCAATCTATAGTCAATTAACCAATCGGGAACAAATGATGAT 1084
 DB 1070 TGCTGAAATTAAGGCAATCTATAGTCAATTAACCAATCGGGAACAAATGATGAT 1129
 QY 1085 GAATCTTTGGCTCCAAACATTTGCTTCAAGAAATCTCCAAAGAACATTTGTTGA 1144
 DB 1130 GAACCTGTGCTCCAAACATTTGCTTCAAGAAATCTCCAAAGAACATTTGTTGA 1189
 QY 1145 TGATATGTCACACAGAAATGCTTCTATAGAGCTGTGAAATCCAGAACATTTGT 1204
 DB 1190 TGACATGTCACACAGAAATGCTTCTATAGAGCTGTGAAATCCAGAACATTTGT 1249
 QY 1205 TGTATTAAGTATGTCACAGAAATGCTTCTATAGAGCTGTGAAATCCAGAACATTTGT 1264
 DB 1250 TGTATTAAGTATGTCACAGAAATGCTTCTATAGAGCTGTGAAATCCAGAACATTTGT 1309
 QY 1265 AGAGATATTCATGAGTGAAGAGCAATTTGTTTGAACAACATGCGAGATTTCCCT 1324
 DB 1310 AGAGATATTCATGAGTGAAGAGCAATTTGTTTGAACAACATGCGAGATTTCCCT 1369
 QY 1335 CTTAAGTGTCTCTATATTTTGAATGCTTCTTCTGCTGAGCTGACATTAAGTGA 1384
 DB 1370 CTTCAGCTGTCTCTATATTTTGAATGCTTCTTCTGCTGAGCTGACATTAAGTGA 1429
 QY 1385 GTTAAAGCTGAAGTGAAGAGAAATTCATCATTTCCACAGTTGCTAACATCTCTG 1444
 DB 1430 GTTAAAGCTGAAGTGAAGAGAAATTCATCATTTCCACAGTTGCTAACATCTCTG 1489
 QY 1445 CTACCTCAACAGGCTCTCTGTGTCACCGGATACACAGTGTGAATGCAATTTGCA 1504
 DB 1490 TTACTTCAACAGGCTCTCTGTGTCACCGGATACACAGTGTGAATGCAATTTGCA 1549

QY 1505 GCAGGCGCAATGCTGAAGAAACATATGAGGCTTGTGTTGATTTGGCCCAAGAGATA 1564
 DB 1550 GCAGGCGCAATGCTGAAGAAACATATGAGGCTTGTGTTGATTTGGCCCAAGAGATA 1609
 QY 1565 CATGATCTTGAAGTACCAAGTGAAGATGGAC 1596
 DB 1610 CATGATCTTGAAGTACCAAGTGAAGATGGAC 1641
 RESULT 3
 US-09-677-064-10
 ; Sequence 10, Application US/09677064
 ; Patent No. 6291224
 ; GENERAL INFORMATION:
 ; APPLICANT: Martino-Calt, Susan J.
 ; APPLICANT: Wang, Hongyu
 ; APPLICANT: Beach, Larry R.
 ; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
 ; TITLE OF INVENTION: Plants and Uses Thereof
 ; FILE REFERENCE: 0706D
 ; CURRENT APPLICATION NUMBER: US/09/677,064
 ; CURRENT FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: 60/055,446
 ; PRIOR FILING DATE: 1997-08-11
 ; PRIOR APPLICATION NUMBER: 60/055,526
 ; PRIOR FILING DATE: 1997-08-08
 ; PRIOR APPLICATION NUMBER: 60/053,944
 ; PRIOR FILING DATE: 1997-07-28
 ; PRIOR APPLICATION NUMBER: 09/118,442
 ; PRIOR FILING DATE: 1998-07-17
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 1931
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-09-677-064-10
 Query Match 55.9%; Score 984; DB 3; Length 1931;
 Best Local Similarity 76.1%; Pred. No. 3, 5e-296;
 Matches 1212; Conservative 0; Mismatches 380; Indels 0; Gaps 0;
 QY 5 TCTTTATTCCTTTGTAATTTTCATTTCTTAAATTTGTAATAATTTGTAATTCATCA 64
 DB 50 TCGTACCTCGCTTGCATTTCTTCAAAAGAGGAGGAAAGCAAGATTTCAATCA 109
 QY 65 GAATTTTAAAGTGTGATGCTTATGTAAGTACACGAGATTCAGTCCGTGA 124
 DB 110 GAGCTTCGCGCTGAGAGGCTCCACGTCGCTGACCGGATTCAGTCCGTGA 169
 QY 125 CAATTAAGAAACCAACGAACTTTGTTACGAGAAACAGAAATGACCTATCAGTGAATGT 184
 DB 170 CCGGTACGACACGAGAGGCTGTACACGAGGCAAGGACGCGCTCACTGAGGTGT 229
 QY 185 CAATCCCAATTTGTCATTAATGAAATTTAAACCAACATTCATTTCTTAATTTAGGGGT 244
 DB 230 CCGCCCAAGTTCCTCAAGTACATTTCCGACCAACCCGCTCCCAAGCTCGGGGT 289
 QY 245 AATGCTTGTGAGTGTGAGGAGGAAACAGGCTCAACCTCAACCGGAGGTTAATTTGTTAA 304
 DB 290 GATGCTTGTGAGGAGGAGGAGGACACAGGCTTCACGCTGACGCTGAGGATTTGCTCA 349
 QY 305 CCGAGAGGCAATTTGATGAGGCTTAAAGAGACAAAGTTCAACAGCAATTTACTTTGGCTC 364
 DB 350 CAGGAGGAGATCTATGAGGCGAACAGAGACAAAGTGAAGCAAGCACTACTACGAGCTC 409
 QY 365 CTTACCCCAAGCTTCAAGTATTCAGGTTGGGCTCTTCCAGGAGAGGAAATTAAGCCCC 424
 DB 410 CTTACCCCAAGCTTCAAGTATTCAGGTTGGGCTCTTCCAGGAGAGGAAATTAAGCCCC 469
 QY 425 ATTCAAGAGCGCTTCAATGGTTAACTGACGCAATTTGTTGGGGATGGATAT 484
 DB 470 GTTCAAGAGCGCTTCCATAGTGAACCGACGCAATTTGTTGGGGATGGATAT 529

```
QY 485 CAGCAACATGAACCTGCTGATGCGCCAGGGCAAGGCTGTTGACATGATTTGA 544
    |||||
Db 530 TAGCAACATGAACCTGCGCGCATGACAGGGCAAGGCTGCGATATGACCTGCA 589
QY 545 GAAGCAGTTGAGGCTTACATGGAATCCATGCTTCACTCCCGAATCATGACCCGA 604
    |||||
Db 590 GAAAGAGCTCAGGCGCTCATGAGATCCATGCTGCACTCCCGATCATGATCCGA 649
QY 605 TTTCATGCTGCAACCAAGAGAGGCTCCAAACAGCATCAAGGGCAACAACAGA 664
    |||||
Db 650 CTTCATCGGGCTAACCAAGGCTCTGCGCCCAAGCTGTCATCAACAGGCAACAGAGA 709
QY 665 GCAGATTCAACAAATCATCAAGACATCAAGCGTTTAAAGAAAGCCACCAAGTGA 724
    |||||
Db 710 ACAGGTGAGCAGATCATCAAGATATCAAGGAGTTTAAAGAGAAAGAAAGTGA 769
QY 725 GGTGTTGATCTGTGAGCTGCCAACAAGAGTACAGTAATTTGTTGGGCTTTAA 784
    |||||
Db 770 GATAGTTGTTGTGAGCTGCAAAACACTGAAAGTATAGCAATGTGTGCGCTGCTCA 829
QY 785 TGACCAATGAGAAATCTTGGCTGCTGTCGACAGAAATGAGGCTGAGATTTCTCTTC 844
    |||||
Db 830 CGACAGAGAGAAATCTAAGCTGCTGTCGACAGAAATGAGGCTGATCAACATC 889
QY 845 CACTTGTATGCTGCTGCTGTTATGGAATGTTCTTTCATTTAATGAAAGCCCTCA 904
    |||||
Db 890 AACACTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949
QY 905 GAAACCTTTTGTACCAAGGCTGATGATCTTCCATCGGAGAAACACTTGTATGTTGG 964
    |||||
Db 950 GAAACCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1009
QY 965 AGATACCTTCAAGAGTGTGACAGCAAAATGAAATCTGCTGCTGCTGCTGCTGCTGCTG 1024
    |||||
Db 1010 TGACCACTTCAAGAGTGTGACAGCAAAATGAAATCTGCTGCTGCTGCTGCTGCTGCT 1069
QY 1025 GGCTGCTATCAAGCAACATCTATAGTCAATTCATCAACATCTGAGAAACATATGTA 1084
    |||||
Db 1070 TGCTGAAATTAAGCCACTCAATGCTGAGCTACAAACCTTGGGAAACAGATGGCAT 1129
QY 1085 GAATCTTTGGCTCCCAAACTTTCCGTTCCAAAGAAATCTCAAGAGCAAGTGTGTA 1144
    |||||
Db 1130 GAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189
QY 1145 TGATATGCTCAACAGAAATGCAATCTGATAGAGCTGCTGTAACATCAAGCATGTGT 1204
    |||||
Db 1190 TGACATGCTGCTGAGCAATGCAATCTGATAGAGCTGCTGCAATCTGATCATGTGT 1249
QY 1205 TGTATTAATGATGCTTACGTAGAGGAGCAGCAAGAGCCATGATGATGATGATG 1264
    |||||
Db 1280 TGCTATCAAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309
QY 1265 AGAGATATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1324
    |||||
Db 1310 AGAGATATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1369
QY 1325 CTAGCTGCTGCTATTTATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1384
    |||||
Db 1370 CCTGCGGCACTATCATCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1429
QY 1385 GTTTAAAGCTGAAAAATGAGGAAAAATTCACATTCACACCGATGCTTACCATCTCG 1444
    |||||
Db 1430 GCTGAAAGCTGAGGAGAGAGCAAAATTCACATCTCTTCCACCGGCTGCTGCAACATCT 1489
QY 1445 CTACCTCAACAAAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1504
    |||||
Db 1490 TTACTTTCACAAAGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1549
QY 1505 GGAGGCTGCAATGCTGAGAAACATATAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1564
    |||||
Db 1550 GCAAGGAGGCTGCTGAGAAACATATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1609
    |||||
```

```
QY 1565 CATGATTTCTGAGTACAAAGTGAAGCATGGGAC 1596
    |||||
Db 1610 CATGATTTCTGAGTACAAAGTGAAGCATGGGAC 1641
    |||||

RESULT 4
US-09-248-796A-3131
; Sequence 3131, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keich Weinlock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3131
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3131

Query Match 21.9%; Score 384.8; DB 4; Length 1578;
Beet Local Similarity 56.0%; Pred. 4e-109;
Matches 812; Conservative 0; Mismatches 602; Indels 36; Gaps 3;

QY 163 ATGCACTTATCAGTATGATGTCACAAACCAATCTGTCAAATGCAATTTAAACCAACA 222
    |||||
Db 113 ATGCAATATGTAATATATATGTCACCAACCTGTAAGTATGTAATTTAAAGTTGAT 172
    |||||
QY 223 TCCATGTTCTTAATTTAGGGTATGCTTGTGGGTGGGTTGGAACCAAGCTTCAACC 282
    |||||
Db 173 TAAAGTCCCTAAGGTTGTTATGTTAGTGGATGAGGAGTAAACAGCTACTAAT 232
    |||||
QY 283 TCACCGGAGTGTATTTGCTAACCAGAGGCACTTTCATGAGGCTTCAAGAGCAAGATT 342
    |||||
Db 233 TGTATGCTGCACTTACAGCAACAAACATATTTCTTTGAAACAAAGAGTGTG 292
    |||||
QY 343 AACCAAGCAATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
    |||||
Db 293 TTAACCAATTAATGATGTTCTGTTACTAAGCATTCATGTAATTCGATGATTA 352
    |||||
QY 397 CTTTCAAGGAGAGGAATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456
    |||||
Db 353 AAGAACTGAGAGATGTTTACGTTCCATTTAATCTCATTGTTCCATGTTTAAACCA 412
    |||||
QY 457 AGCAATGTTGTTGGGAGTGGGATGATGCAACATGAACTGAGTGCATGAGCA 516
    |||||
Db 413 ATGATTTGTTGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 472
    |||||
QY 517 GGGCAAGGTTTGAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
    |||||
Db 473 GAGCTAAAGTCTTGATGTTTCTTAAAGCAAGTGTATCTACTGGAAGAACAG 532
    |||||
QY 577 TTCACTGCTCCCGGAATCTATGACCGGATTTTCACTTGTGCTGCTGCTGCTGCTGCTGCT 636
    |||||
Db 533 AACCTTTGAAATCAATCTACCTGATTTCACTTGTGCTGCTGCTGCTGCTGCTGCTGCT 592
    |||||
QY 637 ACAAGTCATCAAG-----GGCAAAAGCAAGCAAGTTT 672
    |||||
Db 593 ACAATGTTTAAACCAAGTCAATGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 652
    |||||
QY 673 AACCAATCATCAAGATCAAGGCTTTAAGAAAGCAACCAAGTGAAGCAAGTGTG 732
    |||||
Db 653 AAAAATCAAGAAAGATTCAGAGTTTCAAGGCAAGAAAGAAATTAATTAATTAATTA 712
    |||||
QY 733 TACTGTGACTGCAACACAGAGATACAGTAATTTGCTGTTGAGCTTATGACCA 792
    |||||
```

Db	713	TTTTATGAGCTGCTAATACTGAAAGATATGCTGATGTTTGTGCCAATATGTCATGACACTG	772
Qy	793	TGAGAAATCTCTTGGCTGCTGTGGACAGAAATGAGGCTGAGATTTCTCTCCACTTGT	852
Db	773	CTGAACACTTGATCAAGAGATATTAAGATCAACGAGAGAAATTTGCTCCAGTACTGTCT	832
Qy	853	ATGCCATTGCTGTGTGTATGAAAAATGTTCCCTTACATTAATGAGAGCCCTCGAACACTT	912
Db	833	TTGCCGTTGCTTCTATCTTGGAAAAATGTCATATATTAATGTTTCAACACAAACACAT	892
Qy	913	TTGTACACGAGGCTGATGTATCTTGGCATCGCAGGAAACACTTTGATGGTGGAGATGACT	972
Db	893	TTGTTCCCGGAGTTATTTAGAGTTAGCTGAAAAATTAGACTCATTTCAATTGGTGGTATGATT	952
Qy	973	TCAAGAGTGTGCAGACCAAAATGAATCTGTGTGTGATTTCTTGTGGGGGCTGGTA	1032
Db	953	TCAAGTCAAGGTCAAACAAAATTAATCAGTGTAGCTCAATTTCTGGTCGATGCTGGTA	1012
Qy	1033	TCAAGCCCAACCTTATAGCTTCAACCATCTGGGAAACATATGATATGAAATCTTT	1092
Db	1013	TCAACCACTTTCATATGCTTCTTAAATCACTTGGGTAACATATACGGTTTCAATTTAT	1072
Qy	1093	CGGCTCCACAACTTTCGTTCCAGAGAAATCTCCAGAGCAAGCTGTGTGATGATATG	1152
Db	1073	CATCACCAAAACATTTATGATCTAAGAAATTTCCAAACATCTGTTGTATGATATTA	1132
Qy	1153	TCAACAGCAATGCCAT-----CCTCTATGAGCTGTGTGAACATCCAGACCATGTTGTG	1206
Db	1133	TTGATATCCAAAGAAATATTTGATCAACAGAAATCTGTGACAAAGTTGATCACTGTATCG	1192
Qy	1207	TTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGCCATGATATGATCACTTCAAG	1266
Db	1193	TCATTTAAATACCTTGGCAGCTGTGTGTGATTTCTAAAGTTGCCATGATGATCTATTCAG	1252
Qy	1267	AGATATTCATGGGTGGAAGAGCAACATGTTGTTGCACACACATGCGAGATTCCTCT	1322
Db	1253	AATTAATGTTGGGTGTGATCAACACAAATTAATATTCACAAATGTTGTGAAGATCATTCG	1312
Qy	1327	TAGCTGCTCCTATTAATCTTGGACTTGTCTTCTTGTGAGTCAAGCTAGATACGAT	1386
Db	1313	TTGCTACACCAATGATTAATGACTTGGTGGTGGCTACGAAATTTGCACATAGATTTCAAG	1372
Qy	1387	TTAAAGCTGAAAAATGAGGAAAAATTCCTACTATTCACCCAGTTGCTACATCTCAGCT	1446
Db	1373	TCAAGGCTCTGTAATCTGATTAATGATGAAATGTATCTGTGCTGATTTATGTCTT	1432
Qy	1447	ACCTCACCAAGGCTCCTGCTGTCCACGGGGTACACAGTGTGAAATCATTTGCAAGC	1506
Db	1433	ACTGGCTCAAGGCTCTCATTCAGAGACAGAGATTCAAACCTATCAACGATTTAAACAAAC	1492
Qy	1507	AGCGTGCATCTGTGAAAAACATTAATGAGGGCTGTGTGTGAAATGGGCCCCAGAGATTAAC	1566
Db	1493	AAGCTCAACATTTAGTCACTTGTTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1552
Qy	1567	TGATTTCTGA 1576	
Db	1553	TAAAGATTGA 1562	
RESULT 5			
US-09-397-787-34			
Sequence 34, Application US/09397787			
Patent No. 6468758			
GENERAL INFORMATION:			
APPLICANT: Benson, Darin R.			
APPLICANT: Lodes, Michael J.			
APPLICANT: Mitcham, Jennifer L.			
APPLICANT: King, Gordon E.			
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN			
FILE REFERENCE: 210121.466C2			
CURRENT APPLICATION NUMBER: US/09/397.787			
CURRENT FILING DATE: 1999-09-16			

Db 928 GCTCGTGGGGCTCCCGCACAGAACCATGCTCTCGGAACAGAA 974
RESULT 6
US-09-734-237B-72
; Sequence 72, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734, 237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494, 921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 1602
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-72

Query Match 19.3%; Score 340.2; DB 4; Length 1602;
Best Local Similarity 55.4%; Pred. No. 3.4e-95;
Matches 775; Conservative 0; Mismatches 573; Indels 51; Gaps 4;

QY 224 AAATTAGGGGTAATGCTTGTGGGTTGGGGTGAACCAACGCTCAACCTTCACCGGTGT 293
DB 133 AAACATAGGAATATGCTCATTTGGGTTAGGTGGCAACAATGCTCCACTTTAGTGCTTCG 252
QY 294 GTTATTGCTAACCGAGAGGCAATTTCAATGGGCTAACAAAGACAATTAACAGCCAT 353
DB 253 GTATTGGCGAAATAGACAAATGTGAGATTCAAACTTAAGAGAGCGCTTAAGCAACCAAC 312
QY 354 TACTTGGCTCCCTCCACCCAGAGCTACGCTATCCGAGTTGGGTCTCTCC--AGGAGAG 410
DB 313 TACTTCGGCTCCATGACTCAATGTTCTTACCTTGAACCTGGTATGATCGGAGGGGAAT 372
QY 411 GAAATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTAACCTTCAGCAGCATTTGT 470
DB 373 GAGCTTATGCTCTTTTAACTCTGTTGCCATGCTAGCCCAACGACTTTGTGCTC 432
QY 471 GGGGGAATGGGATATGAGCAACATGAACCTGGCTGATGCCATGCCAGGCAAAAGTTT 530
DB 433 TCTGTGGGACATCAATTAACGCAATCTATGAGAGCTATGCAAGAAATCAAGTTCTC 492
QY 531 GACATCGATTTTCAGAAAGAGTTGAGGCTTACATGGAATTCATGCTTCCATCCCGGA 590
DB 493 GAATATGATCTGCAACAGCGCTTGAAGGGAAGATCTCTTGTAAGGCTCTTCTCTCC 552
QY 591 ATCTATGACCCGGAATTTCAATGCTGCAACCAAGAGAGCGTTCACCAACGTCATCA- 649
DB 553 ATTACTACCTCGATTTCAATGCACTAATCAAGATGAGAGGCAATTAATGCAATCAAT 612
QY 650 -----GGGCACAAAGCAAGCAAGTAAACAATCATC 683
DB 613 TTGGATGAAAAAGCAACGTAAACAGAGGGGTAAATGAGCCCATCTGAAAGCAATCA 672
QY 684 AAAGACATCAAGCGCTTTAAGGAAGCCACCAAGTGAAGAGTGTGTAATGCTGAGACT 743
DB 673 CGCGATATCAAAATTTCAAGAGAAACGCCCTTGAATTAAGTAATGTTCTTTGACT 732
QY 744 GCACAACAGAGAGTACAGTAATTTGTTGGGCTTAAATGACACATGAGAAATCTC 803
DB 733 GCAAAATACGAGAGGTACGTAAGATATCTCCGTGTTAATGACACATGAGAAACCTC 792
QY 804 TTGGCTGCTGTGACGAATAGGCTGATTTTCTCTTCCACTTGTATGCTATGTGT 863
DB 793 TTGCAGTCTATTAAAGATGACCATGAAGAGATTGCTCTTCAAGATCTTTGACAGCA 852

QY 864 TGTGTTATGAAAAATGTTCTTTTCAATTAATGAAACCTTCAGAACACTTTTGTACAGGG 923
DB 853 TCTATCTTGAAGGAGTGTCCTCCATATTAATGTTTCAACCGAGAAATACTTTGTTCCCGGC 912
QY 924 CTGATGTAATCTTGGCATGCGAGAGAACCTTGTANTGTGTGAGANTGATTAAGAGTGT 983
DB 913 TTGGTTACGCTGGCTGAGACATGAGGGTACATTCATTTGGGGAGAGATCTCAATCGGGA 972
QY 984 CAGACCAAAATGAATCTGTGTGTTGATTCTTGTGGGGGCTGTATAGCCCAACA 1043
DB 973 CAACCAAGTTGAATCTGTCTTCTGCGCCAGTTCTTAATGATGACAGATATTAACCGGTC 1032
QY 1044 TCTATAGTCACTTACCAACATCTGGAAAACAATGATGATGATCTTTGCGCTCCACA 1103
DB 1033 TCCATGTCATCTTATPACCATTTAGGCAATATGACGGTTATTAATTAATCTGCTCAAAA 1092
QY 1104 ACTTCGCTTCCAGGAATCTCCAGAGCAACGTTGTGATGATGATGATGATGATGAT 1163
DB 1093 CAATTTAGGTCTTAAGGAATTTCCAAAAGTCTGTCAATGATGATGATGATGATGAT 1152
QY 1164 GCCATCTCT-----ATGAGCTGTGAAACATTCAGACCAATGTTGTTATTAATAT 1217
DB 1153 GATATCTTGTACATGATTAATCTGGTAAATAAGTTGACCACTGATGATGATGATGAT 1212
QY 1218 GTGCTTAAGTAGGGGACAGAGAGCCATGATGATGATGATGATGATGATGATGATGAT 1277
DB 1213 ATGAGCCCGTGGGAGCTCAAAAGTGGCAATGACGATATTAACAGTATGATGATGAT 1272
QY 1278 GGTGAAAGAGACCATTTGTTTGCACAAACATGCGAGATTTCCCTTGTAGCTCTCT 1337
DB 1273 GGTGGCCATTAACCGGATTTCAATCAATGTTGCGAAGATTTTACTGGCTACGGCC 1332
QY 1338 ATTATCTGAGCTTGTGCTTCTGAGCTACAGCACTAGATGATGATGATGATGATGAT 1390
DB 1333 TTGATCATGATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1392
QY 1391 -----AGCTGAAATGAGGAAATTCATCTATTCACCCAGTTGCTACATCTCTC 1442
DB 1393 GACCCAGTTAAAGAAATGCTGGCAAAATTCGAAATCTTTATTCAGTTTAACTTCTTG 1452
QY 1443 AGCTACCTACCAAGGCTCTCTGTTCCACCGGTACACAGTGTGTAATGATGATGAT 1502
DB 1453 AGTTACTGTTAAAGCTTCATTAACAAGCCAGATTTTCAACCGGTAAATGCTTAAAC 1512
QY 1503 AAGCAGCTGCAATGCTGAAAAACATATGAGGCTGTGTGATGATGATGATGATGATGAT 1562
DB 1513 AAGCAAAAGCCGCTTGAAAAATTTTAAATGATGATGATGATGATGATGATGATGAT 1572
QY 1563 AACATGATTTCTGAGTACA 1581
DB 1573 GAATTAAGATTGGAAGAGA 1591

RESULT 7
US-09-734-237B-74
; Sequence 74, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734, 237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494, 921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74
; LENGTH: 1605
; TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic gene derived from *Saccharomyces cerevisiae* myo-inositol
 ; OTHER INFORMATION: -1-phosphate synthase, having numerous codons replaced with other
 ; OTHER INFORMATION: a encoding the same amino acids to reduce free energy of folding,
 ; OTHER INFORMATION: and a gly codon inserted after the initiating met codon
 US-09-734-23B-74

Query Match 19.3%; Score 340; DB 4; Length 1605;
 Best Local Similarity 55.5%; Pred. No. 3.9e-95;
 Matches 773; Conservative 0; Mismatches 570; Indels 51; Gaps 4;

234 AAATTAGGGGTAATGCTTGGGTTGGGTAAGAAACAGGCTCAACCCCTCAACGGTGGT 293
 196 AAGCTGGGATATCTGATGATGAGCTGGGTTGTAACAACGCTCACTGCTGATCT 255
 294 GTTATTTCTAACCGAGGCAATTTTCATGGGCTTACAAGACAAGATTCACAGCCAAAT 353
 256 GTTCTGGCAACAAACACAACTGAATTCAGACTTAAGAGAGGTAAACAGCCGAAAC 315
 354 TACTTTGGCTCCCTCAACCCCAAGCTCACTATCCAGTTGGGTCCTCC--AGGAGAG 410
 316 TACTTTGGCTTATGATCTAGTGTCTACTCTGAAGCTGGCAATGATGCTGAAGGTAAC 375
 411 GAAATCTATGCCCATTTCAAGAGCCCTTCCAAATGTTTAACTTCAAGCAATGGTTT 470
 376 GACGTTTACGCTCTGTTCACTCTGCTGCGATGATCTTCCGAACGACTTGGTTT 435
 471 GGGGGATGGGATATCAGCAACATGAACCTGCTGATGATGATGATGATGATGATGAT 530
 436 TCTGGTGGGATATCAACAAACGGGATCTGTAGGAAGCAATGAGGTTCTCAGTTCTG 495
 531 GACATGATTTTGCAGAAAGCAATGAGGCTTACATGAAATTCATGCTTCACTCCCGGA 590
 496 GAATATGATCTGCAACAGGCTGGAAGGCTAAGATCTCTGGTTAAGCACTGCCGTC 555
 591 ATCTATGACCCGGATTTCAATGCTGCCAACAAGAGGCTGCAACAACTCATCAAG 650
 556 ATCTATGACCCGGATTTTCAATGCTGCCAACAAGAGGCTGCAACAACTCATCAAG 615
 651 GGCACAAACAGAGCAAGT-----CAACAACTCATC 683
 616 CTGAGCAAGAAAGGTAAGCTTACTACCCGCTGTAAGTGAATCACTGACGTAACCT 675
 684 AAAGACATCAAGCGCTTAAAGAAAGCCACAAAGTGAAGAGTGTACTGTGACT 743
 676 CCGTATATCAGAACTTCAAGAGGAAAGCGCACTGCAAAAGTTATCTGTAAGT 735
 744 GCAACACAGAGAGTAAGTAATTTGGTGGGCTTAAATGACCAATGAGAAATCTC 803
 736 GCTTAACATGAAAGCTTACGTAAGATATCCCGGGTGAACGATCTATGAAAACTG 795
 804 TTGGCTGCTGGAGCAAGAAATGAGCTGAGATTTCTCTTCACTTGTATGATGCT 863
 796 CTGCAATCTATCAAGAAAGCAACAGGAAATGCTCCCTCAACATCTTCTGCTGTA 855
 864 TGTGTTAGAAAATGCTCTTCAATTAATGAAAGCTTCAAGAACTTTTGTACAGGG 923
 856 TGTATCTGAAAGGCTTACGTAATCAACGCTCTCCCAAGAACTTTTGTACAGGG 915
 924 CTGATTTGATCTTGCATTCGAGGAACATTTGATTTGGTGAATGATCTTCAAGAGTGT 983
 916 CTGGTACAGCTGAGCAACAGAGGTAATCTTATGCTGAGTGAATCTGAATTTGAGC 975
 984 CAGACCAAAATGAAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1043
 976 CAGACTAAATGAAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1035
 1044 TCTATATGATTTACCAACATCTGGAACAAATGATGATGATGATGATGATGATGAT 1103
 1036 TCTATGCTCTTATTAACCACTGGGTTAACAAGAGGCTTACCAACCTGCTGCTGGA 1095
 1104 ACTTTCCGTTCAAGAAATCTTCAAGAGCAAGTGTGATGATGATGATGATGATGAT 1163

1096 CAGTCCGTTCTTAAGAAATCTTAATCTCTGATGAGAGCAATCATCGTTCTAAC 1155
 1164 GCCATCT-----CTATGAGCTGCTGTAACATCCAGACCATGTTGTTTATTAAGTAT 1217
 1156 GACATCTGTACAAACCAAACTGGGTAGAAAGTATGATCACTGATCGTATCAAAATAC 1215
 1218 GTTCCTTACGTAGGGGACAGCAAGAGCAATGATGATGATGATGATGATGATGATGAT 1277
 1216 ATGAAACCGGTTGGTATTTAAAGTTGCTATGAGAGAAATTAATCACTCAAGTATGCTG 1275
 1278 GGTGAAAGAGCAACATTTGTTTGAACAACATGAGAGATTTCCCTTTAGCTGCTCT 1337
 1276 GCGGTCACAAACCTGATCTCTATTCACAAAGTTTGAAGATCTTCTGCTGATACCCG 1335
 1338 ATATCTTGAAGTGTGCTCTTCTGCTGAGCTCAG-----CACTAAGATC 1382
 1336 CTGATCATGACCTGCTGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
 1383 GAGTTTAAAGCTGAATATGAGGAAATTTCACTATTCACACCACTTGTACATCTCTC 1442
 1396 GACCCGTTAAAGAAAGTCTGCAAAATTCGAAACTTCTACCCGTTCTGACCTTCTG 1455
 1443 AGTACCTCAGCAAGGCTCTCTGTTTCCACCGGTTACACAGTGTGATGATGATGATGAT 1502
 1456 TCTTACTGCTGAAGAGCTCTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515
 1503 AAGCAGCGTCAATCTGGAATAATGAGGCTGTGTTGATTTGGCCCCAGAGAAAT 1562
 1516 AAACAGCGTCAATCTGGAATAATGAGGCTGTGTTGATTTGGCCCCAGAGAAAT 1575
 1563 AACATGATTTCTGA 1576
 1576 GAATGCGTTTGA 1589

RESULT 8

US-09-118-442-15

; Sequence 15, Application US/09118442B

; Patent No. 6197561

; GENERAL INFORMATION:

; APPLICANT: Martino-Catt, Susan J.

; APPLICANT: Wang, Hongyu

; APPLICANT: Beach, Larry R.

; APPLICANT: Wang, Xun

; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in

; FILE REFERENCE: 0706

; CURRENT APPLICATION NUMBER: US/09/118,442B

; EARLIER FILING DATE: 1998-07-17

; EARLIER APPLICATION NUMBER: 60/055,446

; EARLIER FILING DATE: 1997-08-11

; EARLIER APPLICATION NUMBER: 60/055,526

; EARLIER FILING DATE: 1997-08-08

; EARLIER APPLICATION NUMBER: 60/053,944

; EARLIER FILING DATE: 1997-07-28

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 15

; LENGTH: 3546

; TYPE: DNA

; ORGANISM: Zea mays

US-09-118-442-15

Query Match 9.1%; Score 160.6; DB 3; Length 3546;
 Best Local Similarity 64.5%; Pred. No. 7e-39;
 Matches 305; Conservative 0; Mismatches 84; Indels 84; Gaps 1;

1032 ATCAAGCAACATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
 2716 ATGAGGCCCACTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2775

```
QY 1092 TCGGCTCCAACTTCCTGCTTCAGAGAAATCTCCAAAGCAAGCTTGTGATGATG 1151
DB 2776 TCTGCCCTTAAACATTTAGAGTCCAGAGAGATCTCCAAAGCAAGCTGTGTGATGATG 2835
QY 1152 GTCAACAGCAATGCCATCTCTATAGAGCTGTGTGAACATCCAGACCATGTGTGTATT 1211
DB 2836 GTCTGAGCAATGCCATCTCTATAGAGCTGTGTGAACATCCAGACCATGTGTGTATT 2895
QY 1212 A----- 1212
DB 2896 AAGGTCTGTAGCTGATCTTTACCTCGTTAAAGTTGACATATGCAAGGACGATTTACA 2955
QY 1213 -----AGTATGCTCTTAAGTGGGAGACAGCAAGAGACC 1247
DB 2956 TTGAACCTTGCTCTTTTGTGAGATGTGCGGTAGCTGGAGAGACAGCAAGAGGCT 3015
QY 1248 ATGATGATGATACCTTTCAGAGATATTTCAATGGGTGAAAGAGACCATTTGTTTGCACAC 1307
DB 3016 ATGAGAGAGATACCTTTCAGAGATCTTTCATGGGCGGCAAGAACACATGCTGTGACAC 3075
QY 1308 ACATGAGAGATTCCTCTTACCTGCTCTATATCTTGAATGACCTTCTTGTGAG 1367
DB 3076 ACCTGTGAGACTGCTCTCTGCGGCACTTATCATCTTGTGATGCTGTGCTGTGAG 3135
QY 1368 CTCAGCACTAGATGAGATTTAAAGCTGAAATGAGGGAATTCACACTCAT 1420
DB 3136 CTCAGCACTAGATGAGATTTAAAGCTGAAATGAGGGAATTCACACTCAT 3188
```

RESULT 9

```
US-09-677-064-15
; Sequence 15, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-15
```

Query Match 9.1%; Score 160.6; DB 3; Length 3546;

Best Local Similarity 64.5%; Pred. No. 7e-39;

Matches 305; Conservative 0; Mismatches 84; Indels 84; Gaps 1;

```
QY 1032 ATCAAGCCAACTCTATAGCTTACCAACCAATCTGGGAAACAAATGATGTGATCTT 1091
DB 2716 ATGCAAGCCAACTCTATAGCTTACCAACCAATCTGGGAAACAAATGATGTGATCTT 2775
QY 1092 TCGGCTCCAACTTCCTGCTTCAGAGAAATCTCCAAAGCAAGCTTGTGATGATG 1151
DB 2776 TCTGCCCTTAAACATTTAGAGTCCAGAGAGATCTCCAAAGCAAGCTGTGTGATGATG 2835
QY 1152 GTCAACAGCAATGCCATCTCTATAGAGCTGTGTGAACATCCAGACCATGTGTGTATT 1211
DB 2836 GTCTGAGCAATGCCATCTCTATAGAGCTGTGTGAACATCCAGACCATGTGTGTATT 2895
```

```
QY 1212 A----- 1212
DB 2896 AAGGTCTGTAGCTGATCTTTACCTCGTTAAAGTTGACATATGCAAGGACGATTTACA 2955
QY 1213 -----AGTATGCTCTTAAGTGGGAGACAGCAAGAGACC 1247
DB 2956 TTGAACCTTGCTCTTTTGTGAGATGTGCGGTAGCTGGAGAGACAGCAAGAGGCT 3015
QY 1248 ATGATGATGATACCTTTCAGAGATATTTCAATGGGTGAAAGAGACCATTTGTTTGCACAC 1307
DB 3016 ATGAGAGAGATACCTTTCAGAGATCTTTCATGGGCGGCAAGAACACATGCTGTGACAC 3075
QY 1308 ACATGAGAGATTCCTCTTACCTGCTCTATATCTTGAATGACCTTCTTGTGAG 1367
DB 3076 ACCTGTGAGACTGCTCTCTGCGGCACTTATCATCTTGTGATGCTGTGCTGTGAG 3135
QY 1368 CTCAGCACTAGATGAGATTTAAAGCTGAAATGAGGGAATTCACACTCAT 1420
DB 3136 CTCAGCACTAGATGAGATTTAAAGCTGAAATGAGGGAATTCACACTCAT 3188
```

RESULT 10

```
US-09-118-442-14
; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-14
```

Query Match 9.0%; Score 159; DB 3; Length 3546;

Best Local Similarity 79.7%; Pred. No. 2.2e-38;

Matches 200; Conservative 0; Mismatches 50; Indels 50; Gaps 1;

```
QY 446 GGTAAACCTGACGACATTTGTTTGGGAGATGGGATATCAGCAATGAACCTTGGCTGA 505
DB 1559 GGTAAACCTGACGACATTTGTTTGGGAGATGGGATATCAGCAATGAACCTTGGCTGA 1618
QY 506 TGCATGCGCAGGAGAAAGTGTGATGATGATTTGCAAGAAAGAGTTGAGGCTTAACT 565
DB 1619 CTCCATGACAGAGGCGCAAGTGTGATGATGATTTGCAAGAAAGAGCTCAGGCTTAACT 1678
QY 566 GGAATTCATGCTTTCAC-TCCCGGAATCTATGACCGGGAATTTGATGCTGCAACAG 624
DB 1679 GGAATTCATGCTTTCACCGGGAATCTATGATGATGATTTGATGCTGCAACAG 1738
QY 625 AGAAGCTGCGCAACAGCTCATCAAGGCAAGAAAGAGCAAGTGAACCAATCATCA 684
DB 1739 GCTTCGCGCAACAGTGTCTATCAAGGCAACCAAGAAAGAGCAAGTGAACCAATCATCA 1798
QY 685 AAGACATCAAG 695
```

Db 1799 AGGATATCAGG 1809

RESULT 11

US-09-677-064-14
; Sequence 14, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-14

Query Match 9.0%; Score 159; DB 3; Length 3546;

Best Local Similarity 79.7%; Pred. No. 2.2e-38; Matches 200; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

Qy 446 GGTTAACCTGAGCATTTGTGTTGGGGATGGATATCAGCAATGAACTGGCTGA 505
Db 1559 GGTGAACCCAGAGCATTTGTGTTGGGGATGGATATGCAATGAACTGGCTGA 1618
Qy 506 TGGCATGGCCAGGCAAGGTTGTTGACATTCATTCAGAGAGCAAGTGGCTTACAT 565
Db 1619 CTCATGACAGGCGCAAGGTGTTGATATGACCTGCAAGACAGCTCAGGCCCTACAT 1678
Qy 566 GGAATCATGCTTCCAC-TCGCCGAAATCTAGAACCCGGAATTCATGCTGCCAACAAG 624
Db 1679 GGATCTCATGTGCTCCACTTCCCGGATATGATCCGGAATTCATGCGCGCTAACCAAG 1738
Qy 625 AGAGGCTGCCAACAAGCATCAAGGCAAGGCAAGCAAGCAAGTTCACCAATCATCA 684
Db 1739 GCTCTGCGCCCAACAGTGTATCAAGGCAAGCAAGCAAGTTCACCAATCATCA 1798
Qy 685 AAGACATCAAG 695
Db 1799 AGGATATCAGG 1809

RESULT 12

US-09-949-016-12608
; Sequence 12608, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12608
; LENGTH: 77626
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(77626)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12608

Query Match 5.1%; Score 89.4; DB 4; Length 77626;

Best Local Similarity 71.8%; Pred. No. 9.3e-16; Matches 117; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1036 AGCCACATCTATAGTCAGTTCACCAACCATCTGGGAACAATGATGTAATCTTCCG 1095
Db 34941 AGACATGTCATCTGATGATTCACCAACCATCTGGCAACAACATGGGGAACCTTACG 35000
Qy 1096 CTCACAACTTCCGTTCCAGAAATCTCCAGAGCAAGTGTGATGATATGTCGA 1155
Db 35001 CGCCATTCAGATTCGCTCTTAAGAGGTGTCCAGAGCAACGTGTGACGACATGTGC 35060
Qy 1156 ACAGCATGCTATCTCTATGAGCTGTGGAACATCTCAACCA 1198
Db 35061 AGAGCAACCCAGTGTCTATACGCCCGGGAAGAGCTTACCA 35103

RESULT 13

US-09-313-294A-4684
; Sequence 4684, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4684
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348858H1
; NAME/KEY: unsure
; LOCATION: 185, 272, 275, 281, 283, 288
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4684

Query Match 3.5%; Score 62.2; DB 4; Length 294;

Best Local Similarity 65.0%; Pred. No. 7e-09; Matches 91; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 54 ATGTTTCATGAGAAATTTTAAGTTGAGTCTCTAATGTGAATACACGAGCTGAGATT 113
Db 125 ATGTTTCATGAGAAAGCTTCGCGCTGAGAGCCCCACAGTGGGTACGGCCGACGAGATC 184
Qy 114 CAGTCGCTGTAACAACAGCAACCAACCACTTGTTCAGAGAAAGAAATGGCACTAT 173
Db 185 NCTCGAGTACCGGTATGACACAGAGAGTGTGTGACAGAGCCAAAGAGCGGCTCC 244
Qy 174 CAGTGGATTTGCAAAACCA 193
Db 245 CGCTGGGTGTCGCCCA 264

RESULT 14
US-09-902-540-5256
; Sequence 5256, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5256
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5256

Query Match 2.8%; Score 48.4; DB 4; Length 1335;
Best Local Similarity 43.4%; Pred. No. 0.0039;
Matches 434; Conservative 0; Mismatches 551; Indels 15; Gaps 4;
QY 427 TCAAGAGCCTGCTTCAATGTTAACCCTGACGACATTTGTTGGGGATGAGATATCA 486
DB 194 TCAAGAGCTGTTGCTCTGCGGAGCTTAAGAGACCTTCCCTTCGGCGCTGGACATCA 253
QY 487 GCACATGAACCTGCTGATGCGATGCGGCGCAAGGTGTTTGAATGATGATTTGCAGA 546
DB 254 TCCGAGAGACCATACGAAGTGGCGGTGCGTGGCGGTGCTCAGCGACAGACCTGG 313
QY 547 AGCAGTTGAGGCTTACATGAGATTCATCTTCCATCCCGGAAATCTATGACCGGAT 606
DB 314 AGAGGTGAAGCGTCTCCAGAGATCAAGCCGAGAGGCGGTGCAGACCTGAT 373
QY 607 TCATTGCTGCCAACAAGAGAGCGTGCACAAAGTCATCAAGGGCAAAAGCAAGAAC 666
DB 374 TCGTGCGCGCATCGAGCGAACACATCAAGGCCACCAAGACGAC-----CGGAGA 427
QY 667 AAGTTCAACAAATCATCAAGAATCATCAAGCGCTTTAAGAGACCAACAAAGTGCACAG 726
DB 428 GCATGAGACGCTGGCGCCAGACATCGGGAATTCAGAGAGAGCTCAAGCGAGCGCG 487
QY 727 TGGTTGATCTGAGATGCCAACACAGAGAGTACATTAATTTGGTGGCTTATG 786
DB 488 CCGTATGCTGTGTGACAGAGCTGAGACCTTCGTCGCTGCCGATCCTTCA-- 544
QY 787 ACACCATGAGATATCTTGGGCTGTGAGACAAATGAGGCTGAGATTTCTCTTCA 846
DB 545 AAGCGTGGCGCTTCGAGAGGCGCTGAGAGAGAGCGCCGACATCAACCCACCG 604
QY 847 CCTTGATGCCATTTGCTGTATGAGAAATGTTCTTTCAATTAATGAGACCTTCA 906
DB 605 CCGTGAACCTTACGCGGCATCAAGAGAGGCGGTGCTTCGGAACGCGCAGCCAAAG 664
QY 907 ACACTTTGTA--CGAGGCTGATGATCTTGGCATCGGAGAGAACATTTGATTTG 963
DB 665 CCAGCGCTGACACGCGCGGCTCCAGAGATGCGCAAGAGAGTGGTGGCGCTGCGG 724
QY 964 GAGATGACTTCAAGAGTGTGACAGACCAAAATGAAATCTGTGTTGTTGATTTCTTGG 1023
DB 725 GCGGCACTTCAAGAGCGCGCCAGACATGATGAGCGGTTCATCGCGCGCGCTTAAG 784
QY 1024 GGGCTGTATCAAGCCCAATCATATAGTCAATTAACCATCTGGGAAACAAATGATGTA 1083
DB 785 CCGGATGCTGGGTGTGATGCTGTCTTCCACCAACATCTGGGCAACCGGAGCGGG 844
QY 1084 TGAATCTTTGGCTCAAACTTTTCGTTCCAGAGAAATCTCCAAAGACAGCTTTTG 1143
DB 845 AGGTGCTGACGACCCCGCGGCTTCAGAGCCAAAGAGTCAACAGTGCAGGCTGTGG 904

QY 1144 ATGATATGTCAACAGCAATGCCATCCTTATGAGCTGTGAAATTCACAGACATTTG 1203
DB 905 AACCATCTCTGACGCC---GACCTGTACCGGAGCTGTACAGAAATGACGGGCAAGAG 961
QY 1204 TTGTTATTAAGTATGCTCTTACCTAGGAGACAGACAGAGCCATGATAGTACTT 1263
DB 962 TGGCCATCTACTACATACCCCGCGCGCGAGCGCAAGAGGTTGGACAAATCGACA 1021
QY 1264 CAGAGATTTTCATGGGTGGAAAGAGACCATTTGTTTCACAAACATGCGAGATTC 1323
DB 1022 TCACCGATGCTGCGCTATCCGATGACGATCAAGTCAATCTTCTGCGGACTTCA 1081
QY 1324 TCTTATGCTCTCTTATTAATCTTGAACCTTGTCTTGTGAGCTCAGACTAGAAATG 1383
DB 1082 TCTGCGCGCGCGCGTGTGCTGAGCATCGGCTGTTCTGAGACTGCGCAAGCGGTGG 1141
QY 1384 AGTTAAAGCTGAATAATGAGGAAATTCACACTATTCCA 1423
DB 1142 AGTGGCGGCGCATCCAGAGATGATGTCTTACTTCA 1181

RESULT 15
US-09-902-540-1257/c
; Sequence 1257, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1257
; LENGTH: 34316
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1257
Query Match 2.8%; Score 48.4; DB 4; Length 34316;
Best Local Similarity 43.4%; Pred. No. 0.0034;
Matches 434; Conservative 0; Mismatches 551; Indels 15; Gaps 4;
QY 427 TCAAGAGCCTGCTTCAATGTTAACCCTGACGACATTTGTTGGGAGATGATATCA 486
DB 19705 TCAAGAGCTGTTGCTCTGCGGAGCTGAAGAGCTTGGCTTCGCGCTGGGACATCA 19646
QY 487 GCACATGAACCTGCTGATGCGATGCGGCGCAAGGTGTTTGAATCATGATTTGCAGA 546
DB 19645 TCCGAGAGACCATACGAAGTGGCGGTGCTCGGGGTGCTCAGACAGACCTTGG 19586
QY 547 AGCAGTTGAGGCTTACATGAGATTCATGCTTCACTCCCGGAATCTATGACCGGAT 606
DB 19585 AGAGGTGAAGCGCTTCTCCAGAGATCAAGCCAGAGGCGCTGACAGACCTTGAGT 19526
QY 607 TCATTGCTGCCAACAAGAGAGCGTGCACAAAGCTCATCAAGGGCAACAAAGCAAGAGC 666
DB 19525 TCGTGCGCGCATTCGAGGCGCAACATCAAGGCCACCAAGCGCAC-----CGCGAGA 19472
QY 667 AAGTTCAACAAATCATCAAGACATCAAGGCTTTAAGAGAGCCCAAAATGAGCAAG 726
DB 19471 GCATGAAGCGGTGCGCAGCATCCGCACTTCAAAAGAGCTCAACGCGAGCGCG 19412
QY 727 TGGTTGATCTGAGCTGCAACAGAGAGGTACAGTAATTTGTTGGGCTTAAATG 786
DB 19411 CCGTATGCTGTGTGACAGAGCGGTGAGACCTTCCGTCGCTGCCAGTCTTCA--- 19355
QY 787 ACACCATGAGAAATCTTGTGCTGTGTGACAGAAATGAGGCTGAGATTTCTCTTCA 846

Db 19354 AGACGCTGGCGCGCTTGAGAGAGCGCGCTGACAGAAACAGCCCGGACATCAACCCACCG 19295
Qy 847 CCTTGATGCGCATGCTTGCTGTTATGAGAAATGTTCCCTTCATTTAATGGAAGCCCTCAGA 906
Db 19294 CGCTGTACACCTTACGCGCCCATCAGAGAGGCGCTTCCGGAACGCCACGCCCAACG 19235
Qy 907 ACACCTTTGTA---CGAGGCGCTGATGATCTTGCCATCGGAGAGAACCTTTGATTTGGTG 963
Db 19234 CAGCGGTGACACGCGCGCGCTTCAGAGAGTGGCCAGAGAGTGGCGGTGGCCGCGCG 19175
Qy 964 GAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTGTGTTGATTTCTTGTGG 1023
Db 19174 GCGGCACTCAAGAGCGGCGCAGACGATGATGAAGAGGTCAATCGCGCCCGCGCTCAGG 19115
Qy 1024 GGGCTGTATCAAGCCAACTTATATGTCAGTTACAACCATCTGGGAAACATGATGTA 1083
Db 19114 CCGGCATGCTGGGTGTGATGCTGGTCTCCACCAACATCTGGGCAACCGCGACGCG 19055
Qy 1084 TGAATCTTCGGCTCCACAACCTTCGTTCCAGAGAAATCTGCAAGAGCAAGCTTGTG 1143
Db 19054 AGGTGCTGACGACCGCGCGCTTCAAGGCAAGAGTCAACCAAGTCGAGGCTGTGG 18995
Qy 1144 ATGATATGCTCAACAGCAATGCCATCTCTATGAGCTGGTGAACATCCAGACCATGTG 1203
Db 18994 ACACCATCTGCGAGCCC---GACTGTACCCGAGCTGTACAGAAAGTACGCGCACAGG 18938
Qy 1204 TTGTTATTTAAGTATGTGCTTACGTAGGGGACAGCAAGAGCCATGATGAGTACACTT 1263
Db 18937 TGGCCATCCACTACTACCGGCCCGGCGGAGCGAAGGAGGTGGACAAACATGACA 18878
Qy 1264 CAGAGTATTCATGGGTGAGAAAGACCACTTTTGGACAAACATGCGAGGATTCCC 1323
Db 18877 TCACCGGATGGCTCGGCTATCCGATCAGATCAAGGTCAACTTCTGCGCGGACTCCA 18818
Qy 1324 TCTTAGCTGCTCTATTTATCTTGGACTTGGTCTTCTGCTGAGCTCAGCACTAGAAATCG 1383
Db 18817 TCCTGGCCGCGCGCTGGTGTGACATGCGCTGTTCCTGAGCTGGCCAGCGGCTGG 18758
Qy 1384 AGTTAAAGCTGAAAAAGAGGAAAAATTCACCTCATTTCA 1423
Db 18757 AGTGGCGGCGCATCCAGAGTGGATGCTTACTTCAA 18718

Search completed: June 8, 2005, 03:03:24
Job time : 312.905 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 22:54:36 ; Search time 5996.19 Seconds

(without alignment)
11172.621 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttattccttctgt.....aatgtaataatttctgt 1760

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hlc:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991.4	56.3	1740	9	CL961092 OaIFCC005
2	819.6	46.6	1828	3	CNS004MK BX827819 Arabidops
3	799.8	45.4	813	7	CK768601 Gm-r1030-
C	4 784.4	44.6	2360	1	AJ583520 AJ583520
5	706.8	40.2	988	7	CK277950 EST724028
C	6 690.8	39.2	782	2	AW348136 GM210001A
7	680.8	38.7	754	4	BI968101 GM830004A
8	658.4	37.4	686	2	AW348857 GM210010A
9	656.4	37.3	936	7	CK279064 EST725142
10	653.4	37.1	914	7	CK275352 EST721430
11	651.8	37.0	672	2	AW508269 s151h10.y
12	647.4	36.8	867	6	CB973407 CAB30002
13	642.6	36.5	865	7	CO085839 GR_Ea03E
14	639.4	36.3	898	7	CO113870 GR_Eb014
15	636.2	36.1	906	7	CK320178 L2P08a02
16	633.2	36.0	891	7	CF212628 CGF100065
17	631.8	35.9	860	7	CO072025 GR_Ba30P
18	631.6	35.9	842	7	CF211061 CAB20007
19	628.8	35.7	836	7	CF518404 CAP0007_I
20	608.4	34.6	648	1	AI794750 sb68f07.y
21	603.4	34.3	860	7	CO082095 GR_Ba46I
22	602.8	34.2	734	6	CA801137 sau01d09.y
23	601.2	34.2	622	2	AW397496 sg79e12.y
24	599.2	34.0	670	4	BG044525 saa29e07.y

25	597.6	34.0	816	6	CB975215
26	595.4	33.8	769	6	CA900545
27	594	33.8	918	4	BG38516
C	28 592	33.6	619	2	BE331363
29	591.6	33.6	762	2	BE660321
30	591.6	33.6	834	7	CK933094
31	591.2	33.6	960	7	CK274458
32	590.8	33.6	746	6	CA918410
33	590.6	33.6	831	7	CO132349
34	586.8	33.3	830	7	CO107630
35	586.2	33.3	909	6	CB345158
36	586	33.3	841	7	CO075472
37	584.6	33.2	618	2	AW279066
38	580.6	33.0	707	2	BF005128
39	579.6	32.9	863	6	CA900542
40	579.6	32.9	663	6	CD438045
41	576.2	32.7	781	6	CB343821
42	573.2	32.6	891	7	CV290142
43	569.2	32.3	897	7	CF514596
44	567.4	32.2	594	4	BM523521
45	566	32.2	575	2	AW397453

ALIGNMENTS

RESULT 1	CL961092	1740 bp	DNA	linear	GSS 21-SEP-2004
LOCUS	OaIFCC005463	Oryza sativa Expressed library	Oryza sativa (indica		
DEFINITION	cultivar-group) genomic, genomic survey sequence.				
ACCESSION	CL961092	GI:52376905			
VERSION	CL961092.1				
KEYWORDS	GSS.				
SOURCE	Oryza sativa (indica cultivar-group)				
ORGANISM	Oryza sativa (indica cultivar-group)				
REFERENCE	1 (bases 1 to 1740)				
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.				
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis				
JOURNAL	Unpublished (2004)				
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.				
FEATURES	Location/Qualifiers				
source	1..1740 /organism="Oryza sativa (indica cultivar-group)" /mol_type="genomic DNA" /db_xref="taxon:39946" /clone_lib="Oryza sativa Expressed library" /note="Oryza sativa exon trapped genomic sequences "				
ORIGIN					
Query Match	56.3% Score 991.4; DB 9; Length 1740;				
Best Local Similarity	78.0%; Pred. No. 1.5e-276;				
Matches 1193; Conservative	0; Mismatches 336; Indels 0; Gaps 0;				
Oy	54 ATGTTTCATGAGAAATTTAAGTTGAGTCTCTTAATGTAAGTACACCGAGCATGATTT 113				
Db	1 ATGTTTCATGAGAAATTTAAGTTGAGTCTCTTAATGTAAGTACACCGAGCATGATTT 60				
Oy	114 CAGTCCGCTGATACACTACGAAACACCGAATCTTGTTCACGAGAACGAGAAATGACACTT 173				

Db 61 GAGTGGAGTACAGATGACAGACGAGGCTGTGTGACAGAGACGACGAGCGGCTCC 120
Qy 174 CAGTGGATTTGCAAAACCCAAATCTGTCAATTCGAATTTAAACCAATCCATTTCT 233
Db 121 CGCTGGGTGTCGGCCGAGTCCGCTGCTACCTTCGGACCAACCAACCTGTCC 180
Qy 234 AAATAGAGGGTATATCTTGTGGGTGTGGGTGGAACAACGCTCAACCTCAACGGTGT 293
Db 181 AAGCTCGGGGTATCTGTGTGGGTGTGGGTGGAACAACGCTCAACCTCAACGGTGT 240
Qy 294 GTTATGTCAACCGAGGAGGCTTTCATGGGCTACAAAGACCAAGATTTCAACAGCTAT 353
Db 241 GTTCATGCGCAACAGGAGGAGGATCTCATGGGCGAACCAAGACCAAGATTTCAACAGCTAT 300
Qy 354 TACTTGTGCTCCCTCAACCCAGGCTCAAGTATCCAGATTTGGTCTTTCAGAGGAAGAA 413
Db 301 TACTATGTGCTCACTACCCAGGCTCAACATCAGGATCGGAGCTTCAACCGGAGAGAG 360
Qy 414 ATCTATGCCCCATTCAGAGGCTGCTTCAATGTGTTAAACCTGACGACATTTGTGTTGG 473
Db 361 ATCTAGGGGCTTCAAGAGCTCTCTGCGCATGTGTGAACCTGTGATGACCTTGTGTGG 420
Qy 474 GGATGGGATATGAGCAACATGAACCTGTGTATGTCATGGCCAGGCGCAAGGTTGTGAC 533
Db 421 GGCTGGGACATTAGCAACATGAACCTGTGTATGTCATGACCAAGGCAAGGTTACTTGAC 480
Qy 534 ATCGATTTGCAAGACAGTTGAGGCTTCAATGAAATCCATGCTTCCACCTCCCGGATTC 593
Db 481 ATGTATCTGCAAGACAGCTCAAGACCTTACATGAGTCCATGTCTCTCCCGGATTC 540
Qy 594 TATGACCCGATTTATGCTCTCCCAACCAAGAGGCTGCAACAGCTCATCTCAAGGAC 653
Db 541 TATGACCCGATTTATGCTCTCCCAACCAAGAGGATTCGCGGCAACATGTCTCAAGGAC 600
Qy 654 ACAAGCAAGACAGATTTCAACAAATCATCAAGATCAAGGCTTTAAGAGCCACC 713
Db 601 ACCAAGAGAGACAGATGAGAGATCATCAAGACATCAAGGATTTCAAGAGAGAGC 660
Qy 714 AAGTGGCAAGGTGTGTACTGTGTGACCTGCAACAGAGATGATGATGATTTGTT 773
Db 661 AAGTGGCAAGGT 720
Qy 774 GTGGGCTTATGACACCATGAGAGATCTCTGGCTGTGTGACAGAAATGAGGCTGAG 833
Db 721 GTTGGGCTTATGACACCATGAGAGATCTCTGGCTGTGTGACAGAAATGAGGCTGAG 780
Qy 834 ATTTCTCTTCCACCTGTGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 893
Db 781 ATATCAACCATGACACTGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 894 GGAAGCCCTCAAGACATTTGTACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 953
Db 841 GGAAGCCCTCAAGACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Qy 954 TTGATTTGGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTGTGT 1013
Db 901 CTGATTTGGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATGATCTGTGTGTGT 960
Qy 1014 TTCTTGTGGGGCTGTGTATCAAGCCAAATGTATAGTCAATTAACCAACCTGTGGAAAC 1073
Db 961 TTCTTGTGGGGCTGTGTATTAAGCCCACTCAATGTGTGTGTGTGTGTGTGTGTGT 1020
Qy 1074 AATGATGTGATGATCTTCCGCTCAACAACTTTCCGTTCCAAGAAATCTTCAAGAGC 1133
Db 1021 AATGATGTGATGATCTTCCGCTCAACAACTTTCCGTTCCAAGAGATCTTCAAGAGC 1080
Qy 1134 AAGTGTGTGATGATGATCTTCAAGAGATGATGATGATGATGATGATGATGATGAT 1193
Db 1081 AATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 1194 GACCATGTTGTGTATTAAGTATGTGCTTTCATGAGGAGACGCAAGAGACCATGAT 1253

Db 1141 GATCATGTTGTGTGATCAAGTATGTGCGGTATGTGAGACAGCAAGAGGCAATGAC 1200
Qy 1254 GAGTACACTTCAGAGATATTCATGAGTGAAGAGACACATTTGTTTTCACAAACATGC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGAGTGAAGAGACACATTTGTTTTCACAAACATGC 1260
Qy 1314 GAGATTCCTCTTGTGCTCTCTTATATCTTGAACCTTGTCTTGTGTGAGCTGAC 1373
Db 1261 GAGACTCACTTCTTGTGCTCTCTTATATCTTGAACCTTGTCTTGTGTGAGCTGAC 1320
Qy 1374 ACTAGAAATCAGTTTAAAGTGAAGAAATGAGGAAATTCACATTCACACCTGCT 1433
Db 1321 ACCAGATTCAGCTGAAGGCGAGGAGAGAGATTCATTCCTTCATTCAGTGTCT 1380
Qy 1434 ACCATCTCAGCTACCTCAGCAAGGCTCTCTGTGTCACCGGATACACAGTGTGAT 1493
Db 1441 GCCCTGGCAAGAGAGGCAATGCTTGAACATCATGAGGCGCTGTGTGTGCTGCTG 1500
Qy 1554 CCAAGAAATACATGATTTCTGAGTACAA 1582
Db 1501 CCGAGAACCAATGATCTGTGAGTACAA 1529

RESULT 2
CNS004MK
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA complete sequence from clone G5LTPG1242E06 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).
BX827819
ACCESSION
BX827819.1 GI:42462602
VERSION
HTC; G5L1_cDNA.
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1828)
REFERENCE
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M. Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL
Unpublished
REFERENCE
Genoscope.
AUTHORS
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)
TITLE
JOURNAL
COMMENT
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction ; Temple G.
Genoscope members carried out sequencing and annotation ; Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Protet_EF/Full_length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
source
Location/Qualifiers
1..1828
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"

/db xref="taxon:3702"
 /clone="GSLR242506"
 /issue="Hormone Treated Callus"
 /plasmid="pCMVSPORT 6"
 complement (1..1828)
 /gene="At4g39800"
 ORIGIN

Query Match 46.6%; Score 819.6; DB 3; Length 1828;
 Best Local Similarity 71.2%; Pred. No. 1.6e-226;
 Matches 1129; Conservative 0; Mismatches 444; Indels 13; Gaps 3;

31 ATTCTTAATCTTGTGAAAAATATATGTTCATCGAGATTTTAAAGTTGATGTCCTAATG 90
 |||||
 35 AATTTTCTTTTCTTCTTCAAAATTTTATGAGCTTCAAGTTGAGAGCCGAAAG 94
 |||||
 91 TGAAGTACACGAGACTGAGATTCAGTCCGTGTACAACTACGAAACACCGAATTGTTG 150
 |||||
 95 TGAAGTACACGAGAGATGAGATTCAGTCCGTGTACAACTACGAAACACCGAATTGTTG 154
 |||||
 151 ACGAGAA--CAGGAATGGCACTTATCAGTGTGTTCAAACTCAATCTTCAATACG 207
 |||||
 155 ATGAGAAAGACCGTCAATGGCACTTACCAATGATGTGAAGCCAAAGACTGTCAATACG 214
 |||||
 208 AATTTAAACCAACATCCATGTTCTTAATTTAGGGTAAAGCTGTGGGTGGGTTGAA 267
 |||||
 215 ATTTCAAAACCGATATTCGTGTCCTCAAAATGAGGGTAAAGCTGTGGGTGGGTTGAA 274
 |||||
 268 ACAACGGCTCAACCTTACCGGTGTGTTATTTGCTAACCGAGAGGCAATTTCAAGGCTA 327
 |||||
 275 ACAATGATCAACTCTCACCGGTGTGTTATTTGCTAACCGAGAGGCAATTTCAAGGCTA 334
 |||||
 328 CAAAGACAAAGTTCAACAAAGCAATTTACTTTGGCTCCCTCAACCGAAGCTCACTATCC 387
 |||||
 335 CCAAGACAAAGTTCAACAAAGCAATTTACTTTGGCTCCCTCAACCGAAGCTCACTATCC 394
 |||||
 388 GAGTTGGGTCTTCCAGGAGAGAAATCTATGCCCCCATTTCAAGAGCTGTCCAAATG 447
 |||||
 395 GTGTGCAATCTTTAAGGTGAAAGATCTATGCTCTTCAAGATCTCTTCCAAATG 454
 |||||
 448 TTAACTTCAGCAATTTGTTGGGGGATGGGATATCAGCAACATGAACCTGCTGATG 507
 |||||
 455 TGAATCCGATGATGTTGTTGGAGATGGGATATTAAGTACATGAATCTTACAGCATG 514
 |||||
 508 CCAATGACAGGCAAGGTGTTGATGATTTGCAAGAGAGTTGAGGCTTACATG 567
 |||||
 515 CCAATGACAGGCAAGGTGTTGATGATTTGCAAGAGAGTTGAGGCTTACATG 574
 |||||
 568 AATCATGCTTCACTCCCGGAATCTATGACCCGATTTTCAATGCTGCCAACAAAGAG 627
 |||||
 575 AAGAACATTTGTCACCTCCCTGGAGATCTGATCTGATTTCAATGCTGCTTAACTCAAGG 634
 |||||
 628 ACGCTGCCAACAAGCTCATCAAGGCAACAAGCAAGAGAGTTCAACAATCATCAAG 687
 |||||
 635 CACGTGCTTACCAAGTATCAAGGCAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 694
 |||||
 668 ACATCAAGGCTTTAAGAGCCACCAAGATGAGCAAGGTGTTGATCTGTGAGCTGCA 747
 |||||
 695 ACATCAAGGCTTTAAGAGCCACCAAGATGAGCAAGGTGTTGATCTGTGAGCTGCA 754
 |||||
 748 ACACAGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAA 807
 |||||
 755 ACACAGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAAAGGAGGCTTAA 814
 |||||
 808 CTGCTGTGACAGAAATGAGGCTGAGATTTCTCTTCACTTGTATGCTTGTGTTG 867
 |||||
 815 ACTGTGTGATGAGGAGTGAAGCTGAGATCTCTCTTCACTTGTATGCTTGTGTTG 874
 |||||
 868 TTATGAGAAATGTTCTCTTCAATTAATGAGAGGCTTCAAGCACTTTGTATCAAGGCTGA 927
 |||||
 875 TTCTTGAAGGAGATCCCGTTCAATTAATGAGAGGCTTCAAGCACTTTGTATCAAGGCTTA 934
 |||||
 928 TTGATCTTGTGCAATCCGAGAGCACTTTGATTTGTTGAGAGTGAATCACTTCAAGAGTGTG 987

Db 935 TTGATATGCGCATCAAGAAATGTTTGTATGCTGCTGAGATATTTCAAGAGTGTGCAAA 994
 |||||
 Qy 988 CCAAAATGAAATCTGTGTTGTTGATTTCTTTGAGGAGGCTGTATCAAGCAATCTTA 1047
 |||||
 Db 995 CCAAGATGAAATCTGTGTTGTTGATTTCTTTGAGGAGGCTGTATCAAGCAATCTTA 1054
 |||||
 Qy 1048 TAGTCAATTAACAATCTGTGGAACCAATGATGATTAATCTTTGAGGCTGTCAAACTT 1107
 |||||
 Db 1055 TTGTAGGCTTACATCACTTGGGAAACAGATGAGATGATTTCTCAAGTTCACAGACTT 1114
 |||||
 Qy 1108 TCCGTTCCAAAGAAATCTTCAAGAGCAAGTGTGATGATTAATGTTCAACAGCAATGCA 1167
 |||||
 Db 1115 TGAATCTTAAGAGATCTTCAAAAGCAATGTGTGAGCATATGTTGTTGATCAATGTA 1174
 |||||
 Qy 1175 TCTCTTTGAGGCTGGGAAACACTGGAC--CCATGTTGTGTCAACAGAGAAACAAG 1233
 |||||
 Db 1228 TAGGGACAGCAAGAGCCATGATGATGATCACTTCAAGATATTTCAAGGTTGAAAGA 1287
 |||||
 Qy 1234 TGAACCAATCAATCAAGAGCAATGAGGAGTTTAAAGAGAAATTAAGTGAATTAAGTTG 1293
 |||||
 Db 1288 GCACCATTTGTTTGCACAAACATGCGAGATTTCCCTTTAGCTGCTCTATTTATCTTGG 1347
 |||||
 Qy 1294 TGGTTCTTGCAGCGCTAACAGAGCTTATGCAATGTGTCTTGGAGTGAACGATA 1353
 |||||
 Db 1348 ACTTGGCTCTTCTTGTGAGTCTGAGCTAGCAATGAGTTTAA-----GCTGAAA 1398
 |||||
 Qy 1354 CAATGAGAAATCTCAATGAGTGTGATGAGGATGAGGCTGATCTCTTCAACAC 1413
 |||||
 Db 1399 ATGAGGAAATTTCAATCTTCCAGCCAGTTGCTAACCTCTCAAGTCTTCAACAG 1458
 |||||
 Qy 1414 TTATGAGAAATTTCAATCTTCCAGCTTGTGAGGCAATCACTCAAGTCTTCAACAG 1473
 |||||
 Db 1459 CTCTCTGCTTCCAGCGGTAACCAAGTGTGATGATGATTTCAAGAGGCTGTGAAATG 1518
 |||||
 Qy 1474 CACCTCTTGTACCGCGGAACTCCGATGATTAAGAACTGTGAGAGAGGCTATTC 1533
 |||||
 Db 1519 TGAAGAAATTAATGAGGCTTGTGATTTGAGGCTTCAAGAAATTAATGATTTCTGAGT 1578
 |||||
 Qy 1534 TGAAGAAATTAATGAGGCTTGTGATTTGAGGCTTCAAGAAATTAATGATTTCTGAGT 1593
 |||||
 Db 1579 ACAATGAAAGATGAGGACCGAAGAT 1604
 |||||
 Qy 1594 TCAAGTGAACATGAGAGCAATGAT 1619
 |||||

RESULT 3
 CK768601 813 bp mRNA linear EST 20-FEB-2004
 LOCUS Gm-r1030-621 Gm-r1030 Glycine max cDNA clone Gm-r1030-621 5', mRNA
 DEFINITION
 ACCESSION CK768601
 VERSION CK768601.1 GI:42722702
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE 1 (bases 1 to 813)
 AUTHORS Perlapuram,C.C., Li,L., Wurtele,E.S., Westgate,M.E. and
 Nikolau,B.J.
 TITLE Molecular determination of soybean composition
 JOURNAL Unpublished (2004)
 CONTACT Basil J. Nikolau and Cyril C. Perlapuram
 Department of Biochemistry, Biophysics and Molecular Biology
 Iowa State University
 2210 Molecular Biology Building, Ames, IA 50011, USA
 Tel: 515 294 9423
 Fax: 515 294 0453

Email: dimas@iastate.edu
 Individual base call and confidence value were assigned using the
 phred software (http://www.phrap.org/). Overall sequence quality
 assessment and vector trimming were conducted using the Lucy
 software (http://www.tigr.org/software/).
 This clone was originally generated by the Public Soybean EST
 Project (http://129.186.26.94/soybeanest.html)/Shoemaker, R
 (rcshoe@iastate.edu).
 This clone is available through: Biogenetic Services, 801 32nd Ave,
 Brookings, SD 57006. For further information call 605-697-8500 or
 contact info@biogeneticservices.com
 Seq primer: 17-1 (5' AAT ACG ACT CAC TAT AG 3').

FEATURES

source

1. 813
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="Gm-r1030-621"
 /lab_host="DH10B"
 /clone_lib="Gm-r1030"
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
 cDNA library was constructed from mRNA isolated from
 immature cotyledons of greenhouse grown plants
 (individual seed fresh weight of 100-300mg). The library
 was prepared using the Life Technologies pSuperScript cDNA
 library construction kit. Complementary DNA was
 synthesized from mRNA using a poly(dT) sequence with a
 NotI restriction site. SalI linkers adapters were ligated
 to the blunt-ended cDNA fragments followed by NotI
 digestion. The cDNA fragments were directionally cloned
 into the NotI-SalI restriction site of the pSPORT1
 vector. The ligated cDNA fragments were transformed into
 E. coli Electromax DH10B host cells. This library was
 constructed by Dr. Ilya Vodkin and Dr. Anu Khanna. Note
 that Gm-r1030 is a re-rack of Gm-cl007."

ORIGIN

Query Match 45.4%; Score 799.8; DB 7; Length 813;

Best Local Similarity 99.8%; Pred. No. 7.1e-221;

Matches 801; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 485 CAGCAACATGAACCTGCTGATGCCCGCAAGGAGGCTTTGACATGATTTGCA 544
DB 11 CGGCAACATGAACCTGCTGATGCCCGCAAGGAGGCTTTGACATGATTTGCA 70
QY 545 GAAGCAGTTGAGGCTTACATGAATTCATGCTTCACCTCCCGAATCATGACCCGA 604
DB 71 GAAGCAGTTGAGGCTTACATGAATTCATGCTTCACCTCCCGAATCATGACCCGA 130
QY 605 TTTCATTCGTCGCAACCAAGAGAGGCTGCAACCAACGTCATCAAGGCAACCAAGA 664
DB 131 TTTCATTCGTCGCAACCAAGAGAGGCTGCAACCAACGTCATCAAGGCAACCAAGA 190
QY 665 GCAAGTTCAACAATCATCAAGAGAGGCTTTAAGAGCCACCAAGTGGACAA 724
DB 191 GCAAGTTCAACAATCATCAAGAGAGGCTTTAAGAGAGCCACCAAGTGGACAA 250
QY 725 GGTGGTTGACGTGGAGCTGCAACAGAGAGTACAGTATTTGGTTGGGCTTAA 784
DB 251 GGTGGTTGACGTGGAGCTGCAACAGAGAGTACAGTATTTGGTTGGGCTTAA 310
QY 785 TGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAGATTTCTCTTC 844
DB 311 TGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAGATTTCTCTTC 370
QY 845 CACCTTTGATGCAATGCTTGTGTTATGGAATATGTTCTTTCATTAATGAAAGCCCTCA 904
DB 371 CACCTTTGATGCAATGCTTGTGTTATGGAATATGTTCTTTCATTAATGAAAGCCCTCA 430
QY 905 GAACACTTTTGAACAGGCGTGAATGATTTGCAATCGGAGAGAAACATTTATATGGG 964
DB 431 GAACACTTTTGAACAGGCGTGAATGATTTGCAATCGGAGAGAAACATTTATATGGG 490

```

```

QY 965 AGATCACTTCAAGAGTGTGACAGCAACCAATGAATCTGTGTGTTGATTTCTTGGG 1024
DB 491 AGATCACTTCAAGAGTGTGACAGCAACCAATGAATCTGTGTGTTGATTTCTTGGG 550
QY 1025 GGCTGGTATCAAGCAATATATAGTACGTTACACATCGGGAACCAATGATGAT 1084
DB 551 GGCTGGTATCAAGCAATATATAGTACGTTACACATCGGGAACCAATGATGAT 610
QY 1085 GAATCTTGGGCTCCACAACTTTCCGTTCCAGAGAAATCTCCAAGCAAGTGTGTA 1144
DB 611 GAATCTTGGGCTCCACAACTTTCCGTTCCAGAGAAATCTCCAAGCAAGTGTGTA 670
QY 1145 TGATATGTCACACAGCAATGCCATCTTATAGGCTGTGTAACATCCAGCATGTTGT 1204
DB 671 TGATATGTCACACAGCAATGCCATCTTATAGGCTGTGTAACATCCAGCATGTTGT 730
QY 1205 TGTATTATAGTATGCTTACGTATAGGAGCAAGACCAATGATGATGATCTTC 1264
DB 731 TGTATTATAGTATGCTTACGTATAGGAGCAAGACCAATGATGATGATCTTC 790
QY 1265 AGAGATATTCATGGGTGGAAGA 1287
DB 791 AGAGATATTCATGGGTGGAAGA 813

```

RESULT 4

AJ583520/c 2360 bp mRNA linear EST 08-JAN-2004

LOCUS AJ583520 PpProcl Physcomitrella patens cDNA clone pp01015066

DEFINITION similar to putative myoinositol-1-phosphate synthase, mRNA

ACCESSION AJ583520.1 GI:40781590

VERSION AJ583520

KEYWORDS EST.

SOURCE Physcomitrella patens

ORGANISM Physcomitrella patens

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;

Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

REFERENCE 1 (bases 1 to 2360)

AUTHORS Hone, A., Egener, T., Luecht, J.M., Holloer, H., Reinhard, C., Schween, G.

and Reek, R.

AN improved and highly standardised transformation procedure allows

efficient production of single and multiple targeted gene-knockouts

in a moss, Physcomitrella patens

Curr. Genet. 44 (6), 339-347 (2004)

CONTACT: Schween G

Plant Biotechnology

University of Freiburg

Sonnenstrasse 5, D-70104 Freiburg, Germany.

location/Qualifiers

1. 2360

/organism="Physcomitrella patens"

/mol_type="mRNA"

/db_xref="taxon:3218"

/clone="pp01015066"

/tissue_type="protonema"

/clone_lib="PpProcl1"

ORIGIN

Query Match 44.6%; Score 784.4; DB 1; Length 2360;

Best Local Similarity 70.4%; Pred. No. 3.1e-216;

Matches 1111; Conservative 0; Mismatches 451; Indels 16; Gaps 4;

```

QY 23 TTTCATTCATCTTAATCTTTGTGAAATATATGTTCAATGAGATTTAAGTTGAGTG 82
DB 2181 TCTCTTGTGTTTATGTTTGTGTGATCAAAATGTTGTCACAGCTTTCTGTGAGAG 2122
QY 83 TCTTATATGAGATACACGAGACTGATTCAGTCCGTGTACACTACGAAACACCA 142
DB 2121 TCTTATATGAGATACACGAGACTGATTCAGTCCGTGTACACTACGAAACACCA 2062
QY 143 ACTTGTACAGAG-----AACAGAAATGCACTTATAGTATGTTCAACCAATC 196

```

Db 2061 GTTCTGACAGACGCGCAAGAGCGATGAGTAGAGTGGTAGCCAAAGCCAAAGTC 2002
 Qy 197 TGTCAATATAGAAATTTAAACCAACATCCATGTTCTTAATTAGGGGTAATGCTTGTGGG 256
 Db 2001 CGTGCCATACCAATTTGGCCATCAGCGCGCAAGGAGGCCAAAGCTGGGAATGTCGTGGG 1942
 Qy 257 TTGGGGTGGAAACAGCGGCTCAACCCCTCAACGGGGGTGTAATGCTAACGAGAGGGCAT 316
 Db 1941 TTGGGGTGGCAACAGGGGTCTCATCTCAGCGCTGGATCTCTGCCCAACAAAGAGTAT 1882
 Qy 317 TTCAATGGGCTACAAAGGA-CAGATTTCAACAAAGCAATTAATTGGCTCCCTCAACCAAG 375
 Db 1881 CTCATGGGAGACCAAGAGATGGCGGTGAGACACCGCAATTACTTGGATCCCTCACTCAAG 1822
 Qy 376 CCTCAGCTATCCGAGTTGGGTCTTCC-----AGGAGAGGAATCTATGCCCATTT 427
 Db 1821 CCTCAGCTATCCGCGGTGGGCTCTCTTCCAGGGTGGAGGAAGATGATGATACCTTTTAA 1762
 Qy 428 CAAGAGCCTGCTTCAATGGTTAAACCTGACGACATTTGTTGGGGGATGGGATATACAG 487
 Db 1761 GAAGCACTTCTAACCGATGGTCAACCCCAACGATCTGATGATGGCGGTTGGGATATCTC 1702
 Qy 488 CAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGAGTGTGATGATCGATTGGAGAA 547
 Db 1701 CAACATGAACCTGGCGGATGCCATGGATGGTCCCGGATGCTGGACATGATCTACAGAA 1642
 Qy 548 GCAAGTGGAGCCTTCAATGGAATTCATGCTTCACTCCCGGAATCTATGACCCGGAATT 607
 Db 1641 GCAACTTGTCTCTCATGAGATATGTATGTTACCGCTGATCTACAGCCCGCACTT 1582
 Qy 608 CATTTGTCGAACCAAGAGAGGCGTGCACAAAGTATCAAGGGGCAACAAAGCAAGAGCA 667
 Db 1581 CATGCTGCTAACCAAGAGAGCGTGTGCAACATCTCATCAAGGGGTTCCAAAGAGAGCA 1522
 Qy 668 AGTTCAACAATCATCAAGAGACATCAAGCGCTTAAAGAGCCCAAAAGTGACAGAGT 727
 Db 1521 AATGAGCGGATATATCCAGAGACATACAGGAATTCAGAGGGCAACCAAGGTTGACAGAGT 1462
 Qy 728 GGTGTATCTGTGATCTGCCAACAAGAGAGTCAAGTAATTTGGTGGGCTTATATGA 787
 Db 1461 AGTGTACTGTGACCGCTAACACAGAGCGTTACAGCGATGATGATGATGATGATGA 1402
 Qy 788 CA-CCATGAGAGATCTTTGGCGTGTGAGACAGAAATGAGCGTGAATTTCTCTTCCA 846
 Db 1401 CACCAAGAGAGATCTTTGATTTCTGTGAGAGAGAGAAATTCCTCCGCTCA 1342
 Qy 847 CCTTGTATCCATGCTTGTGTATGAGAAATGTTCTTTCATTATGAGAGCCCTCAGA 906
 Db 1341 CCTTGTACGCTCTGCTTGTGATCCAGAGAGGTTCTTTCATTCAACGAGTCCCAAG 1282
 Qy 907 ACACTTTGTACAGAGGCTGATGATCTTGCATGCGAGAGAACACTTGTGATGTTGAG 966
 Db 1281 ACACATTTCTCCCTGAGATATGAAATACGCTGTGAAAGAACTATTAATTTGGTGTG 1222
 Qy 967 ATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTTGTGATTTCTTGTGGGGG 1026
 Db 1221 ACACATTTCAAGAGCGGTCAAAAGAGATGAAATCTTCTAGTGAATTTGATGAG 1162
 Qy 1027 CTGTGTATCAAGCAACATCTATAGTATGATTAACAACATCTGGGAAACAATGATGATGA 1086
 Db 1161 CGGGAATTAAGCAAGTCAATTTGTATGCTACAAACATTTGGGCAACAGATGAGATGA 1102
 Qy 1087 ATCTTTGGCTTCACAACTTTCTGTTCCAGAGAAATCTTCAAGAGCAAGCTTGTGATG 1146
 Db 1101 ACTTGTCCCGCTCAAAACATCTGTTCCAGAGAGATCTTCAAGAGCAATGTAGTGAAG 1042
 Qy 1147 ATATGCTCAACAGCAATGCAATCTATAGAGCTGTGTAACAATCAAGCAATGTTGTG 1206
 Db 1041 ACATGAGTGTGCAAAATGATTTATAGCGCGCGGTGAACATCCGACCAAGTATATG 982
 Qy 1207 TTAATTAATGATGTGCTTATCGTAGGGAGACGAAGAGCAATGATGATCACTTCAAG 1266
 Db 981 TGATCAAGTACGTCTTATGTTGGTGAACGAGAGGGCAATGACGATACACTTCCG 922

Qy 1267 AGATATTCAATGGGTGGAAGAGACCATTTGTTTGACAAACACAGCGAGATTTCCCTCT 1326
 Db 921 AGATCTTCAATGGCGGAGCGATCAACATGCTCATGACAAATACGTGCAAGATTTCTTAC 862
 Qy 1327 TAGCTGCTCTTATATCTTGAAGTGTGCTTCTTGTGAGCTCAGCACTAGAAATGAGT 1386
 Db 861 TGGCGCGCGCTCATCTTGGACTTGTGCTTGTGCTGCTGAGCTGCAAGAAATCCAT 802
 Qy 1387 TTAAGCTGAAAATGAGGAAAATTCACACTCAATCCAGCTGCTTACCATCTCAGCT 1446
 Db 801 TCAGAGAGACGAGAGAGAGTTCATTTTCCACCGGTTGCGTACTCTGAGCT 742
 Qy 1447 ACTTCAACAAAGCTCTCTGTGTTCCACCGGATACACCAATGATGATGATGTAAGC 1506
 Db 741 ACTTCAACAAAGCTCTCTGTGTTACCCCAAGTACCCAGTGTGTTAAGGTTGGCTAAC 682
 Qy 1507 AGGTCATATGCTGGAACCATATATAGAGGCTTGTGTTGATTTGCCCCAGAGAAATACA 1566
 Db 681 AGAGGCGATGCTAGAGAAATATCATGCGCGCTGATGGGTTGTCCCGACATATACA 622
 Qy 1567 TGATTTCTGAGTACAGT 1584
 Db 621 TGATGCTAGATATCAAGT 604
 RESULT 5
 CK277950 988 bp mRNA linear EST 03-AUG-2004
 EST724028 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAE302 5' end, mRNA sequence.
 CK277950
 CK277950.1 GI:39834928
 EST.
 Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
 asterids: lamiales: Solanales: Solanaceae: Solanum.
 1 (bases 1 to 988)
 Buell, C.R., Hart, A., Zisemann, V., Karamycheva, S.A. and Baker, B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST724029
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: ATG TAG GTG ACA CTA TAG.
 Location/Qualifiers
 1..988
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAE302"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="vector: pCMVSPORT1; Site 1: BclRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation
 of the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d

Retracking was performed by Genome Systems, St. Louis,
<http://www.genomesystems.com>, and sequencing by the Keck
 Center for Comparative and Functional Genomics,
 University of Illinois,
<http://www.life.uiuc.edu/bioelect/keck.html>.

ORIGIN

Query Match 39.2%; Score 690.8; DB 2; Length 782;
 Best Local Similarity 92.1%; Pred. No. 3,8e-189;
 Matches 721; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

QY 968 TCAGCTCAAGAGTGTGACAGCAAAATGAATGTGTTGATTTCTTGGGGG 1027
 DB 782 TAGANNNNNNAGNNNTCAGACCAANNNGAAATCTGTTGGTTGATTTCTTGGGGG 723
 QY 1028 TGGTATCAAGCCCAATCTATAGTCAAGTTCAACCACTGGGAAACATGATGATGA 1087
 DB 722 TGGTATCAAGCCCAATCTATAGTCAAGTTCAACCACTGGGAAACATGATGATGA 663
 QY 1088 TCGTTGGCTCCCAAACTTCCGTTCCAGGAAATCTCCAGAGCAAGCTTGTATGA 1147
 DB 662 TCCNCGGCTCCCAAAANNNTCCGCTCCAGGAAATCTCCAGAGCAAGCTTGTATGA 603
 QY 1148 TATGCTCAAGCAATGCAATCTCTATAGAGCTGGTGAACATCCAGCATGTTGT 1207
 DB 602 TATGCTCAAGCAATGCAATCTCTATAGAGCTGGTGAACATCCAGCATGTTGT 543
 QY 1208 TATTAAGTATGTCCTTACGTAGGAGGAGCAAGCAAGAGCCATGATGATGATGAT 1267
 DB 542 TATTAAGTATGTCCTTACGTAGGAGGAGTACAGAGCAAGCATGATGATGATGATGAT 483
 QY 1268 GATATTCATAGGCTGGAAAGAGCACTTGTGTCACACATGCGAGATTCCTCTT 1327
 DB 482 GATATTCATAGGCTGGAAAGAGCACTTGTGTCACACATGCGAGATTCCTCTT 423
 QY 1328 AACTGCTCTATATATCTTGGACTTGTGCTGTCGAGCTGAGCACTAGATGAGTT 1387
 DB 422 AACTGCTCTATATATCTTGGACTTGTGCTGTCGAGCTGAGCACTAGATGAGTT 363
 QY 1388 TAAAGCTGAAATGAGGAGGAAATTCACATCTTCCAGCTTGTGATCTCTAGCTA 1447
 DB 362 TAAAGCTGAAATGAGGAGGAAATTCACATCTTCCAGCTTGTGATCTCTAGCTA 303
 QY 1448 CCTCAAGAGGCTCTCTGTTCCACCGGGTACACCAAGGTGATGATGATGATGATGAT 1507
 DB 302 TCTGACCAAGGCTCTCTGTTCCACCGGGTACACCAAGGTGATGATGATGATGATGAT 243
 QY 1508 GCGTGCATGCTGAAACATATATGAGGGCTGTTGATGAGGCTGAGGAGATTAACAT 1567
 DB 242 GCGTGCATGCTGAAACATATATGAGGGCTGTTGATGAGGCTGAGGAGATTAACAT 183
 QY 1568 GATTCGAGTACAGAGTGAAGCAATGAGGAGCAAGAAATATATATGTTGGGATGAGCT 1627
 DB 182 GATTCGAGTACAGAGTGAAGCAATGAGGAGCAAGAAATATATATGTTGGGATGAGCT 123
 QY 1628 GATGCTTTATATATATATATATGTTGCTTATATATTTTCCAAAGTGAATGATGATGA 1687
 DB 122 GATGCTTTATATATATATATATGTTGCTTATATATTTTCCAAAGTGAATGATGATGATGAT 63
 QY 1688 GCTTCATTAATGCTTATAGAGGAGGAGCAATGCTGTTTATAGAGCAATGAATGATGAT 1747
 DB 62 TC-TCAATTAATGCTTATAGAGGAGGAGCAATGCTGTTTATAGAGCAATGAATGATGATGAT 4
 QY 1748 TAT 1750
 DB 3 TAT 1

RESULT 7
 B1968101/c B1968101 754 bp mRNA linear EST 23-OCT-2001
 LOCUS GM830004A22D05 Gm-r1083 Glycine max cDNA clone Gm-r1083-1306 3',
 DEFINITION mRNA sequence.

ACCESSION B1968101
 VERSION B1968101.1 GI:16342506
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

REFERENCE
 AUTHORS Vokkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
 Erpelding, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L., and Lewin, H.
 TITLE A Functional Genomics Program for Soybean (NSF 9872565)
 JOURNAL Unpublished (1999)
 COMMENT Other ESTs: A1855435 corresponding to Gm-c1013-777 (5')
 Contact: Vokkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics

University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vokkin@uiuc.edu
 This clone is available through: Incyte Genomics, 4633 World
 Parkway Circle St. Louis, Missouri 63134, Phone (800) 430-0030 or
 (314) 427-3222 FAX: (314) 427-3324. Web site:
<http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboratio>
 n/index
 Seq primer: 5'-TTTTTTTTTTTTTTTTTTT (A/C/G) -3'.

FEATURES

source
 1..754
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone_id="Gm-r1083-1306"
 /clone_id="Gm-r1083"
 /note="The library Gm-r1083 is a sequence-driven, rerecked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-c1009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-c1013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-c1028 (from 'Superpod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were rerecked to form library Gm-r1083. The cDNA clones of the rerecked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, <http://web.anc.umn.edu/biodata/netsoy/>. Retracking was performed by Incyte Genomics, St. Louis, <http://www.incyte.com>, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.life.uiuc.edu/bioelect/keck.html>. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 38.7%; Score 680.8; DB 4; Length 754;
 Best Local Similarity 94.4%; Pred. No. 3e-186;
 Matches 711; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 991 AATGAATCTGTGTTGTTGATTTCTTGGGGGCTGATCAAGCAACATCTATAG 1050


```
Db      754  AANTAAATCTGTGTTGGTGGANNNNNTTGGGGGGCTGGATCAAGCCNNNATCTTAAG 695
Qy      1051  TCAGTTACAAACCATCTGGGAAACATGATGATGAATCTTTGGCTCCACAACTTCC 1110
Db      654  TTAGTTACACACATCTGGGAAACATGATGATGAATCTCTGGCTCCACAACTTCC 635
Qy      1111  GTTCCAGGAAMTCTCCAGAGCAACGTTTGTGATGATATGGTCAACAGCAATGCCATCC 1170
Db      634  GCTCCAGGAAMTCTCCAGAGCAACGTTTGTGATGATGATGATGATGATGATGATGATG 575
Qy      1111  TCTATGAGCCTGCTGTAACATCCAGACATGTTTGTGATGATGATGATGATGATGATG 1230
Db      574  TCTATGAGCCTGCTGTAACATCCAGACATGTTTGTGATGATGATGATGATGATGATG 515
Qy      1231  GGGACAGCAAGAGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATG 1290
Db      514  GGGATAGCAAGAGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATG 455
Qy      1231  CCATTGTTTGGCAACACATGCGAGAGATTCCTCTTACCTGCTCTTATATCTTGGACT 1350
Db      454  CCATTGTTTGGCAACACATGCGAGAGATTCCTCTTACCTGCTCTTATATCTTGGACT 395
Qy      1351  TGGTCTCTTCTGCTGAGCTCAGACCTAAGATGATGATGATGATGATGATGATGATG 1410
Db      394  TGGTCTCTTCTGCTGAGCTCAGACCTAAGATGATGATGATGATGATGATGATGATG 335
Qy      1411  TCCACTCATTCACACCAAGTGTGCTACATCTCAGCTACCTCAGCAAGGCTCCTCTGTT 1470
Db      334  TCCACTCATTCACACCAAGTGTGCTACATCTCAGCTACCTCAGCAAGGCTCCTCTGTT 275
Qy      1471  CACCGGGTACACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1530
Db      274  CACCGGGTACACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 215
Qy      1531  TGAGGGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1590
Db      214  TGAGGGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 155
Qy      1591  TGGGACCGAAGATATATATATGATGATGATGATGATGATGATGATGATGATGATG 1650
Db      154  TGGGATAGAAATATATATATGATGATGATGATGATGATGATGATGATGATGATG 95
Qy      1651  TTTGCTTATATATTTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1710
Db      94  TTTGCTTATATATTTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 36
Qy      1711  GGCATATTTCTGTTACTAGGAACATGATGATGATGATGATGATGATGATGATGATG 1743
Db      35  GACATATTTCTGTTTCTAGGAGACATGATGATGATGATGATGATGATGATGATGATG 3

RESULT 8
AM348857/c 686 bp mRNA linear EST 04-OCT-2000
LOCUS      GM210010A10E12 Gm-r1021 Glycine max cDNA clone Gm-r1021-3671 3',
DEFINITION mRNA sequence.
ACCESSION  AM348857
VERSION     AM348857.1 GI:6846567
KEYWORDS   EST
SOURCE      Glycine max (soybean)
ORGANISM   Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1 (bases 1 to 686)
            Vodka, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
            Erdling, J., Rapp, C., Shoop, B., Pardinas, J., Liu, L. and Lewin, H.,
            A Functional Genomics Program for Soybean (NSF 9872565)
            Unpublished (1999)
            Other ESTs: A1496426
            Contact: Vodka, L.O., PI, A Functional Genomics Program for
```

```
Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional
Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodka@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or info@genome
systems.com web site: www.genomesystems.com
Seq primer: 5'-TTTATTATTTTATTTT(A/C/G)-3'.

FEATURES
    source
        1..686
            /organism="Glycine max"
            /mol_type="mRNA"
            /cultiivar="Williams"
            /db_xref="taxon:3847"
            /clone="Gm-r1021-3671"
            /issue_type="root"
            /lab_host="XLA0-G01d"
            /clone_lib="Gm-r1021"
            /note="Vector: Bluescript II XR; Site 1: EcoRI; Site 2:
            XhoI; Library Gm-r1021 is a sequence-driven, retracted set
            of the original library Gm-cl004 which was prepared from
            root cDNA. The mRNA was isolated from entire roots of 8
            day old 'Williams' seedlings which were propagated on
            paper towels with distilled water. Stratagene's cDNA
            Synthesis Kit (catalog #200401) was used to synthesize
            the cDNA. The Gm-cl004 library was constructed by Dr.
            Paul Keim & Virginia H. Coryell, Department of Biology,
            Box5640, Northern Arizona University, Flagstaff, AZ
            86011, email: paul.keim@uau.edu, virginia.coryell@uau.edu.
            The contig analysis is to select unique genes was performed
            by the Laboratory of Ernest Retzel, Computational Biology
            Center, University of Minnesota,
            http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html
            . Rerecting was performed by Genome Systems, St. Louis,
            http://www.genomesystems.com, and sequencing by the Keck
            Center for Comparative and Functional Genomics,
            University of Illinois,
            http://www.life.uiuc.edu/biotech/keck.html."
```

ORIGIN

Query Match 37.4%; Score 658.4; DB 2; Length 686;
Best Local Similarity 96.1%; Pred. No. 9.7e-180;
Matches 659; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```
Qy      1048  TAGTCAGTTACACCATCTGGGAAACATGATGATGATGATGATGATGATGATGATGATG 1107
Db      686  TAGTCAGTTNNAACCAANNNNNNNNNATGANNATGATGATGATGATGATGATGATGATG 627
Qy      1108  TCCGTTCCAGGAAMTCTCCAGAGCAACGTTGTTGATGATGATGATGATGATGATGATG 1167
Db      626  TCCGTTCCNNNNNATCTCCAGAGCAACGNNGTGATGATGATGATGATGATGATGATGCA 567
Qy      1168  TCCCTATGAGCCTGCTGTAACATCCAGACATGTTGTTGATGATGATGATGATGATGATG 1227
Db      566  TNNCTATGAGCCTGCTGTAACATCCAGACATGTTGTTGATGATGATGATGATGATGATG 507
Qy      1228  TAGGGACAGCAAGAGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATG 1287
Db      506  TAGGGAGAGCAAGAGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATG 447
Qy      1288  GCACCATGTTTGGCAACACATGCGAGATTCCTCTTACCTGCTCTTATATCTTGG 1347
Db      446  GCACCATGTTTGGCAACACATGCGAGATTCCTCTTACCTGCTCTTATATCTTGG 387
Qy      1348  ACTGTCCTCTCTTCTGAGAGCTCAGACCTAGAAATGAGTTAAAGCTGAAATGAGGAA 1407
Db      386  ACTGTCCTCTCTTCTGAGAGCTCAGACCTAGAAATGAGTTAAAGCTGAAATGAGGAA 327
```

QY	1408	AATTCATCTCAATTCACCCAGATGGTACATCTCAGCGTACCTCAGCAAGGCTCCTCG	146
Db	326	AATTCATCTCAATTCACCCAGATGGTACATCTCAGCGTACCTCAGCAAGGCTCCTCG	267
QY	1468	TTCCACCGGGTACACCACTGGTGAATGCAATGTCTCAAGCAGCGTCAATGCTGAAAA	1527
Db	266	TTCCACCGGGTACACCACTGGTGAATGCAATGTCTCAAGCAGCGTCAATGCTGAAAA	207
QY	1588	GCATGGGACCGAAGATATATATAGTTGGGGTAGCCTTACGTGAATGTTTATATATAT	1647
Db	146	GCAATGGGACCGAAGATATATATAGTTGGGGTAGCCTTACGTGAATGTTTATATATAT	87
QY	1648	ATGTTGGTATATATTTTGGCAAGTGAATGCAATGCAATGCAATGCAATGCTTAAAG	1707
Db	86	ATGTTGGTATATATTTTGGCAAGTGAATGCAATGCAATGCAATGCAATGCTTAAAG	27
QY	1708	CGGGGCAATATCTGTTTCTAGGAAC	1733
Db	26	CGGGGCAATATCTGTTTCTAGGAAC	1
RESULT 9			
LOCUS	CK279064		
DEFINITION	CK279064	936 bp mRNA linear EST 03-AUG-2004	
ACCESSION	EST725142	potato abiotic stress cDNA library Solanum tuberosum cDNA	
VERSION	CK279064	5' end, mRNA sequence.	
KEYWORDS	CK279064.1	GI:39836042	
SOURCE	EST		
ORGANISM	Solanum tuberosum (potato)		
	Solanum tuberosum		
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	Asterids; lamids; Solanales; Solanaceae; Solanum.		
REFERENCE	1 (bases 1 to 936)		
AUTHORS	Buell,C.R., Harr,A., Zisemann,V., Karanmycheva,S.A. and Baker,B.		
TITLE	Generation of ESTs from abiotic stressed potato tissue		
JOURNAL	Unpublished (2003)		
COMMENT	Other ESTs: EST725143		
	Contact: Robin Buell		
	The Institute for Genomic Research		
	9712 Medical Center Dr, Rockville, MD 20850, USA		
	Email: potato-array@tigr.org		
	Clones can be requested from the University of Arizona Genomics		
	Institute via http://genome.arizona.edu/orders/		
	Seq primer: ATT TAG GTG ACA CTA TAG.		
FEATURES	location/qualifiers		
SOURCE	1..936		
	/organism="Solanum tuberosum"		
	/mol_type="mRNA"		
	/cultivar="Kennebec"		
	/db_xref="taxon:4113"		
	/clone="POAE989"		
	/tissue_type="abiotic stress treated leaf and root tissue"		
	/lab_host="DH10B-Tona"		
	/clone_id="potato abiotic stress cDNA library"		
	/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;		
	supplier: Solanum tuberosum var. Kennebec plants were		
	grown from cuttings on a 1hr light/8 hr dark cycle at 25		
	C for 3-4 weeks. Abiotic stress conditions were applied to		
	four separate sets of plants. Set 1 involved saturation of		
	the soil with 150 mM NaCl and tissues were harvested at		
	following application of the salt stress (leaves: 2hr,		
	6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).		
	Set 2 were grown under the standard conditions and then		
	were water stressed by withdrawal of further watering		
	applications. Drought stressed plants were harvested after		
	cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d		

and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 °C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d and 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 °C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

Query Match		37.3%;	Score 656.4;	DB 7;	Length 936;	
Best Local Similarity		82.1%;	Pred. No. 4,1e-179;			
Matches	768;	Conservative	0;	Mismatches	166;	Indels 2; Gaps 1
Qy	538	ATTTCGAGAGACAGTGTAGGCTCTTACATGGAATCATGCTCTTCACTCCCGGAATCTATG	597			
Db	1	ATTCGCAAAAGACAGCTAGAGGCGCTTACATGGAATCATGCTCTTCTCTGTATCTATG	60			
Qy	598	ACCCGGAATTCATTTGCTGCCAACCAAGAGGAGCGTGCCAAACAAGTCAATCAAGGCAACA	657			
Db	61	ACCTTGACTTCATTGCTAGCTTACCAAGAGACGTTGCCAACAAGTATCAAAAGAACCA	120			
Qy	658	AGCAAGACAGTTCACAAATCATCAAGAATCAAGGCGTTTAAAGAACCAACCAAG	717			
Db	121	AGAAAGAACAGTTGATCAATTTGTAAAGTATTTAGGAATTCACAGGAAGAACAAAG	180			
Qy	718	TGACACAGAGGTGTGTACTGTGACGTGCCAACACAGAGAGGTACAGTATTTTGTTGG	777			
Db	181	TAGCAAGATAGTGTTCTATGACCTGCCAACACCGAAAGGTACAGCATATGTGTGTTG	240			
Qy	778	GCCCTAATGACACCATAGGAATCTCTTGCGCTGTGACAGAAATAGGCTGAGATT	837			
Db	241	GCCTTAACGATACCATAGGAACCTTTTACGTGCTGTGATAGAAATAGGCTGAATAT	300			
Qy	838	CTCCTTCACCTGTGATGCCATTGCTGTGTATGAGAAATGTTCTTTCAATTAATGAA	897			
Db	301	CTCCTTCATCAATGATGATGCTATTGCTGTATTCGAGAAATGTCTTTTCAATCAACGAA	360			
Qy	898	GCCTCGACACACTTTTGTACAGAGGCGTGAATGATCTTGCCATGCGACAGAAACCTTGA	957			
Db	361	GCCTCGAAACACTTTTGTTCAGAGCTTATGTATGATTTGGCCATTAAGGAAACACTTTAA	420			
Qy	958	TTGTGTAGAGATGACTTCAGAGGTGTGACACCAAAATGAATCTGTGTGTTGATTTCC	1017			
Db	421	TTGTGTGTATGACTTTAAGAGGTGTCAACCAAGATGAAGTCTGTGTGTTGATTTCC	480			
Qy	1018	TTGTGTGGGGCTGTGTATCAAGCCAACTATATGCTAGTTACACCAATCTGGGAAACAATG	1077			
Db	481	TTGTGTGAGGTGTATTAAGCCCAACGTCAATAGAGCTCAAACTCTGGGTAAACATG	540			
Qy	1078	ATGATATGATCTTTGGGCTCCCAAACTTTCCGTTCCAGAGAAATCTCCAAAGACAAG	1137			
Db	541	ACGGAATGATCTTTCTGCTCTCCCAACCTTCGGGTCTAAGAGATCTCAAAAAGTATG	600			
Qy	1138	TTGTGTATGATATGTGTCAACAGCAATGCCATCTCTATGAGCTGTGTAAATCTCAAGC	1197			
Db	601	TTGTGTATGATCAATGTGTGTGCAAACTCCATTTCTATGATCTGGGAGACCTTGACC	660			
Qy	1198	ATGTGTGTATTAATGATATGTGCTCTTACGTAAAGGGGACACAAAGAGAGCATGTAGT	1257			
Db	661	ATGTCTGTGTATCAAGTATGTTCCATATGTTGAGACAGACAAAGGCAATGTATGAT	720			
Qy	1258	ACACTTCAGAGATATTCATGGGTGAGAAAGACACCATGTGTTTTCACAACAATGCAAG	1317			
Db	721	ACATGTCAGAGATTTTCAATGGGCGGAAAGACACATAGTTATGACAAACACTTGTGAGG	780			
Qy	1318	ATTTCCTCTTAAAGTGTCTCTTATATCTTGACATTTGTGCTCTTCTGAGCTAGACATA	1377			
Db	781	ACTCTCTTTTGGAGGCTCCAAATATCTTGACATTTGTGCTCTTCTGAGCTAGACACTC	840			

REFERENCE AUTHORS

JOURNAL
COMMENT

FEATURES
Source

REFERENCE	TITLE	COMMENT
1 (bases 1 to 672)	Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rietter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R./Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@waterston.wustl.edu
	This clone is available through: Reggen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cdu@reggen.com	Insert Length: 1864 Std Error: 0.00 High quality sequence stop: 431.
FEATURES	location/Qualifiers	
source	1..672	
	/organism="Glycine max"	
	/mol_type="mRNA"	
	/db_xref="taxon:3847"	
	/clone="GENOME SYSTEMS CLONE ID: Gm-r1030-2492"	
	/lab_host="DH10B"	
	/clone_lib="Gm-r1030"	
	/note="vector: pSPORT1, Site 1: SalI, Site 2: NotI. This cDNA library was constructed from mRNA isolated from immature cotyledons of greenhouse grown plants (individual seed fresh weight of 100-300mg). The library was prepared using the Life Technologies psuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restriction site. SalI linkers adaptors were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note that Gm-r1030 is a re-rack of Gm-cl007."	
ORIGIN		
Query Match	37.0%; Score 651.8; DB 2; Length 672;	
Best Local Similarity	98.1%; Pred. No. 8e-178;	
Matches 659; Conservative 0; Mismatches 13; Indels 0; Gaps 0;		
18 TGTAAATTCATCTCAATCTTATCTTTGGAAAAAATTAATGTCATCGAGAAATTTAAAGTT 77		
1 TGTAAATTCATCTCAATCTTATCTTTGGAAAAAATTAATGTCATCGAGAAATTTAAAGTT 60		
78 GAGTGTCTATATGGAAGTACACCGAGACTGAGATTAGTCCGTGTACAACTACGAAACC 137		
61 GAGGTCTCTATATGGAAGTACACCGAGACTGAGATTAGTCCGTGTACAACTACGAAACC 120		
138 ACCGAACTGTTCACGAGAACGGAATGGACCTATACAGTGGATTTGCAACCCAAATCT 197		
121 ACCGAACTGTTCACGAGAACGGAATGGACCTATACAGTGGATTTGCAACCCAAATCT 180		
198 GTCAATATACGAATTTAAACCAACATCCATCTTCTAAATTAAGGGGTAATGCTGTGGGT 257		
181 GTCAATATACGAATTTAAACCAACATCCATCTTCTAAATTAAGGGGTAATGCTGTGGGT 240		
258 TGGGGTGGAAACAACGGCTCAACCCCTCAACGGTGGTATTATGTCAACCGAGAGGGCATTT 317		
241 TGGGGTGGAAACAACGGCTCAACCCCTCAACGGTGGTATTATGTCAACCGAGAGGGCATTT 300		
318 TCATGGGCTACAAAGACAGATTCAACAGCCAAATTAATCTTGGCTCCCTCACCCCAAGCC 377		

Db	301	TCATGGGCTACAAAGACAAAGATTCAACAAAGCCAAATTACTTTGGCTCCTCACCCAAAGCC	360
QY	378	TCAGCTATCCGAGTTGGGTCCTTCACAGGAGAGAAATCTATGCCCATTCAGAGGCTG	437
Db	361	TCAGCTATCCGAGTTGGGTCCTTCACAGGAGAGAAATCTATGCCCATTCAGAGGCTG	420
QY	438	CTTCACATGGTTAACCCCTGACGACATTTGTTGGGGGATGGGATATACCAACATGTAAC	497
Db	421	CTTCACATGGTTAACCCCTGACGACATTTGTTGGGGGATGGGATATACCAACATGTAAC	480
QY	498	CTGCGCTATGCGCATGCGCCAGGCGAAAGGTTTGGACATCGATTGCGAAGCAGTTGAG	557
Db	481	CTGCGCTATGCGCATGCGCCAGGCGAAAGGTTTGGACATCGATTGCGAAGCAGTTGAG	540
QY	558	CCTTACATGGAAATCGATGCTTCCACTCCCGGAATCTATGACCCTGGATTTCAATTGCTGC	617
Db	541	CCTTACATGGAAATCGATGCTTCCACTCCCGGAATCTATGACCCTGGATTTCAATTGCTGC	600
QY	618	AACCAAGAGAGCGTGCACCAACAGTATCTAAGGCGCAAAAGCAAGCAGTTTCAACAA	677
Db	601	AACCAAGAGAGCGTGCACCAACAGTATCTAAGGCGCAAAAGCAAGCAGTTTCAACAA	660
QY	678	ATCATCAAGAC 689	
Db	661	ATCCTTCAAGAC 672	
RESULT 12			
LOCUS	CB973407	867 bp	mRNA linear EST 01-MAY-2003
DEFINITION	CBAB0002_IC_Fc_D10 Cabernet Sauvignon Berry Stage I - CAB3 Vitis		
ACCESSION	CB973407		
VERSION	CB973407.1	GI:30296613	
KEYWORDS	EST.		
SOURCE	Vitis vinifera		
ORGANISM	Vitis vinifera		
REFERENCE	Ekaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.		
AUTHORS	1 (bases 1 to 867)		
TITLE	Goes da Silva,F., Iandolo,A., Lim,H., Baek,J., Lee,J., Xu,J., Jones,K. and Cook,D.		
JOURNAL	Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages		
COMMENT	Unpublished (2003)		
	Contact: Douglas Cook, PhD		
	CABs Genome Facility		
	UC Davis, Plant Pathology		
	One Shields Ave, Davis, CA 95616, USA		
	Tel: 530 754 6561		
	Fax: 530 754 6617		
	Email: drcoc@ucdavis.edu		
FEATURES	Seq primer: ACGGTACCGACATATGCC.		
source	Location/Qualifiers		
	1..867		
	/organism="Vitis vinifera"		
	/mol_type="mRNA"		
	/cultivar="Cabernet Sauvignon"		
	/db_xref="taxon:29760"		
	/clone="CAB30002_IC_Fc_D10"		
	/sex="hermaphrodite"		
	/dev_stage="Berry stage I"		
	/lab_host="DH5alpha"		
	/clone_id="Cabernet Sauvignon Berry Stage I - CAB3"		
	/note="Organ: Berry; Vector: pDNR; Site_1: SfiI; Site_2: SfiI; CAB3 is a cDNA library of Vitis vinifera 'Cabernet Sauvignon' Clone 8 berries. Samples were collected after berry set from field-grown vines during stage I of berry growth, 17 days after full bloom. The average berry size was 6 millimeters. Sampled vines were located at the University of California, Davis, Experimental Vineyard."		

CDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGTATCAAGCAGAGTGGCATTAAGCGCG-3' and 5'-ATTCTAGAGCGCGGCGGCGGCGATG-dT(30)-NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 36.8%; Score 647.4; DB 6; Length 867;
Best Local Similarity 84.2%; Pred. No. 1.7e-176;
Matches 729; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```
QY 108 GAGATTCACTCCGTGTACACTACGAAACCCGAACTTGTTCACGAGAACGAAATGGC 167
DB 1 GAGATTCACTCCGTGTACACTACGAAACCCGAACTTGTTCACGAGAACGAAATGGC 60
QY 168 ACCTATCAGTGGATGTGTCAACCCAAATGTCAAATGCAATTTAAACCAACATCCAT 227
DB 61 ACCTATCAGTGGATGTGTCAACCCAAATGTCAAATGCAATTTCAAGCTGATGCCAT 120
QY 228 GTTCCCTAAATTTAGGGGTATATGCTTGTGGGTGGGAGAAACAGGCTCAACCTCACC 287
DB 121 GTCCCAATCTAGGGGTATATGCTTGTGGGTGGGAGAAACAGGCTCATTCTCACT 180
QY 288 GGTGGTGTATTTGCTTAACCGAGAGGSCATTTCAATGGCTTCAAGAGACAGATTCACAA 347
DB 181 GGTGGTGTATTTGCTTAACCGAGAGGGAATTCATGGGCAACAGAGACAGATTCACAA 240
QY 348 GCATTAATCTTGGGCTCCCTCAACCAAGCTCAGCATTCGAGTGGGCTTTCACAGGA 407
DB 241 GCATTAATCTTGGGCTCCCTCAACCAAGCTCAGCATTCGAGTGGGCTTTCACAGGA 300
QY 408 GAGAAATCTATGCCCATTTCAAGACCTGCTTCCAAATGGTTAAACCTTGAACGACATTTG 467
DB 301 GAGAGATTTAAGCTCCATTTCAAGAGCTCCCTCCATGGTGAACCCAGATGACATTTGT 360
QY 468 TTTGGGGGATGGGATATCAACCAACCTGCTGATGCATGGCCAGGCAAAAGTG 527
DB 361 TTTGGGGGATGGGATATCAACCAACCTGCTGATGCATGGCCAGGCAAAAGTG 420
QY 528 TTTGATCATGATTTTGAGAGAGAGTTGAGGCTTCAATGGAATCCATGCTTCACTCCC 587
DB 421 CTGACATTTGATCTGAGAGAGAGTTGAGGCTTCAATGGAATCCATGCTTCACTCCC 480
QY 588 GGAATCTATGACCCGATTTCAATGCTGCAACCAAGAGAGCGTGCACCAACGTCATC 647
DB 481 GGAATCTATGACCCGATTTCAATGCTGCAACCAAGAGAGCGTGCACCAACGTCATC 540
QY 648 AAGGCAACAAGCAAGAGAGAGTTCAAAATCATCAAAAGACATCAAGGCTTTAAGAA 707
DB 541 AAGGCAACAAGCAAGAGAGAGTTCAAAATCATCAAAAGAGTTTAAAGAG 600
QY 708 GGCACCAAGTGGAGCAAGGTGTTGATCTGAGACCTGCCAACAAGAGAGTACGTAT 767
DB 601 GGCACCAAGTGGAGCAAGGTGTTGATCTGAGACCTGCCAACAAGAGAGTACGTAT 660
QY 768 TTGTTTGGGGCTTATATGACACATGAGAAATCTTGGCTGCTGTGAGACGAATAG 827
DB 661 TTGTTTGGGGCTTATATGACACATGAGAAATCTTGGCTGCTGTGAGAGATAG 720
QY 828 GCTGAGATTTCTCTTCCACTTTGATGTCATGCTTGTGTTATGAAAAATGTTCCCTTC 887
DB 721 CCTGAGATTTCTCTTCCACTTTGATGTCATGCTTGTGTTATGAAAAATGTTCCCTTC 780
QY 888 ATTATGGAAGCCCTCAGAACACTTTTGAACAAGGCTGATGATCTTCCATCCGAGG 947
DB 781 ATTATGGAAGCCCTCAGAACACTTTTGTTCAGAGGCTGATGATTTGCTATTAAGAG 840
QY 948 AACACTTTGATTTGTTGAGATGACTT 973
DB 841 AACACTTTGATTTGTTGAGATGATTT 866
```

RESULT 13
CO085839 865 bp mRNA linear EST 16-JUN-2004
LOCUS GR_Ea03B19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea03B19 5',
DEFINITION mRNA sequence.

ACCESSION CO085839
VERSION CO085839.1 GI:48776473
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 865)
Kim,H.,Yu,Y.,Kudrna,D.,Hatfield,D.,Stum,D.,Mueller,C.,
Udall,J.A.,Rapp,R.A.,Wendel,J.F.,Rao,K.,Soderlund,C. and

AUTHORS

Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 03 row: E column: 19.

FEATURES

source location/Qualifiers
1..865
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea03B19"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by Agt. More glycerol clones held in -80."

ORIGIN

Query Match 36.5%; Score 642.6; DB 7; Length 865;
Best Local Similarity 83.9%; Pred. No. 4.2e-175;
Matches 726; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```
QY 479 GGATTCAGCAATGAACCTGGCTGATGCCATGGCCAGGCAAAAGTGTGACATGCA 538
DB 1 GGACATTTGATGACATGAACCTGAGTCTATGGCTAAGGCAAGTTTCGACATGGA 60
QY 539 TTTGAGAAAGCAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATCTATGA 598
DB 61 TCTGCAAAAGCACTGAACCTTACATGGAATCCATGATCCCATCTCCGTAATTCAGA 120
QY 599 CCCGATTTCAATGCTGCCAACAAGAGAGCGTGCACCAACGTCATCAAGGGCACAA 658
DB 121 TCTGATTTCAATGCTGCCAACAAGAGAGCGTGCACCAATATGTCATCAAGGGCACAA 180
QY 659 GCAAGAGCAAGTTCAACAATCATCAAAAGACATCAAGCCGTTTAAAGAGCAACCAAGT 718
DB 181 GAAAGAACAAAGTTCCAGCGTGCATCAAAAGACATCAAGAGATTCAAAAGAAAAACAAAGT 240
QY 719 GGACAGGTTGTTGATCTGAGACCTGCCAACAAGAGAGTACAGTAATTTGTTGTGG 778
DB 241 GGACAGGTTGTTGATCTGAGACCTGCCAACAAGAGAGTACAGTAATTTGTTGTGG 300
QY 779 CTTTATGACACATGAGAAATCTTGGCTGCTGTGAGCAAGAAATGAGGCTGAGATTC 838
DB 301 GCTAATATGACACCGTGGAAAGCTTATAGCTTCTTTTGGAGAAAGATGAAATCAGAGATTC 360
QY 839 TCTTTCACCTTGTATGTCATGCTTGTGTTATGAAAAATGTTCTTTCATTAATGAAAG 898
```

Db 361 TCCTTCACCTGTTATGATGATGCTGTTGTTCTTGAAGAAAGTTCCTTCATCATGAGCG 420
QY 899 CCTTCAGAACCTTTTGTACACGAGGCTGATTTGATCTTCCATCGCAGAGAACCTTTGAT 958
Db 421 CCCACAAAACACCTTTGTTCCAGGGTTGATTTGGCTATTCAGAGAACTGTCTGAT 480
QY 959 TGTGAGATGATCTTCAAGATGAGTGCAGCCAAATGAATCTGCTTGGTATTTCT 1018
Db 481 TGAAGAGATGATCTTCAAGATGAGTGCAGCCAAATGAATCTGCTTGGTATTTCT 540
QY 1019 TGTGAGGAGCTGTATCAAGCCAAATCTATGATGATGATCAACATCTGGGAAACATGA 1078
Db 541 TGTGGGAGCTGTGATCAAGCCAAATCTATGATGATGATCAACATCTGGGAAATATGA 600
QY 1079 TGTGATGATCTTTGGGCTTCACAACTTTCCGTTCCAAAGAAATCTCCAGAGCAAGT 1138
Db 601 TGGCATGATCTGTCCAGACCCCAACCTTCCGTTCCAAAGAGATCTCAAGAGCATGT 660
QY 1139 TGTGATGATGATGATCAAGCAATGCTATCTATGAGCTGTGAAATCTCCAGAGCA 1198
Db 661 TGTGATGATGATGATCAAGCAATGCTATCTATGAGCTGTGAAATCTCTGATCA 720
QY 1199 TGTGTTGTTATTAAGATGATGCTTACGTAGGGGACAGCAAGAGCCATGATGATGA 1258
Db 721 TGTGTTGTTATCAAGTATGTTGATGATGATGATGATGATGATGATGATGATGATGA 780
QY 1259 CACTTCAGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 1318
Db 781 CACATCAGAGATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGA 840
QY 1319 TTCCCTTAGCGCTCTATTTATC 1343
Db 841 TTCCCTGTTGCTGCTCCATTTATC 865

RESULT 14
LOCUS COL13870 898 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_Eb014121.r GR_Eb Gossypium raimondii cDNA clone GR_Eb014121
3', mRNA sequence.
ACCESSION COL13870.1 GI:48812557
VERSION COL13870.1
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 898)
Kim H., Yu Y., Kudrna D., Hatfield J., Scum D., Mueller C.,
Wing R.A., Rapp R.A., Wendel J.F., Rao K., Soderlund C. and
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 014 row: I column: 21.
Location/Qualifiers
1..898
FEATURES
Source
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb014121"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_1lb="GR_Eb"
/note="Vector: pCMV.SORT-6.1; Site_1: NotI; Site_2:

ECORY: Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

Query Match 36.3%; Score 639.4; DB 7; Length 898;
Best Local Similarity 82.0%; Pred. No. 3.6e-174;
Matches 726; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY 198 GTCAATATGAAATTTAAACCAATTCATGTTCTTAATTAAGGGATATGCTTGAGCT 257
Db 1 GTCAATATGAAATTTAAAGCTGATACCTATGCCCTTAATTAAGGGATATGCTTGAGCA 60
QY 258 TGGGTTGAAAACAAAGGCTCAACCTTCACCGGTGTTATTTGCTTACCAAGAGGCAATT 317
Db 61 TGGGAGGAAAACAAATGTTCAACCTTCACCGGTGTTATTTGCTTACCAAGAGGCAATT 120
QY 318 TCATGGGCTCAAAAGGACAAATTCATCAAGGCAATTAATTTGGCTCCCTCAAGGCGC 377
Db 121 TCTTGGGCTCAAAAGGACAAATTCATCAAGGCAATTAATTTGGCTCCCTCAAGGCGC 180
QY 378 TCAGCTATCCGAGTTGGGTCTTCCAGAGGAGAAATCTATGCCCCATTCAGAGCGTG 437
Db 181 TCAGAGATCCGAGTTGGGTCTTCCAGAGGAGAAATCTATGCCCCATTCAGAGCGTG 240
QY 438 CTTCCAAATGTTTAACTTCAGAGCAATGTTTGGGGATGAGATATCAAGCAATGAC 497
Db 241 CTTCCAAATGTTTAACTTCAGAGCAATGTTTGGGGATGAGATATCAAGCAATGAC 300
QY 498 CTGGCTGATGCGGATGCGGCAAGGAGGTTTGAATCATGATTTGAGAAAGAGTTGAGG 557
Db 301 CTAGCTGATGCGGATGCGGCAAGGAGGTTTGAATCATGATTTGAGAAAGAGTTGAGG 360
QY 558 CCTTACATGGAATTCATGCTTCCACTCCCGGAATCTATGACCGGATTTCAATTCGTCG 617
Db 361 CCTTACATGGAATTCATGCTTCCACTCCCGGAATCTATGACCGGATTTCAATTCGTCG 420
QY 618 AACCAAGAGAGCGTCCCAACAGCTCATCAAGGCGCAAGAGCAAGCAAGTTCACAA 677
Db 421 AACCAAGAGAGCGTCCCAACAGCTCATCAAGGCGCGCAAGAGCAAGTTCACAA 480
QY 678 ATCATCAAGAGAGCGTCCCAACAGCTCATCAAGGCGCGCAAGAGCAAGTTCACAA 737
Db 481 ATCATCAAGAGAGCGTCCCAACAGCTCATCAAGGCGCGCAAGAGCAAGTTCACAA 540
QY 738 TGAATGCGCAACAGAGAGTACAGTATTTGTTGGGCTTAAATGACCAATGAG 797
Db 541 TGAATGCGCAACAGAGAGTACAGTATTTGTTGGGCTTAAATGACCAATGAG 600
QY 798 AATCTCTGCTGCTGCTGAGCAAGATGAGCTGATTTCTCTTCCACTTGTATGCC 857
Db 601 AATCTCTGCTGCTGCTGAGCAAGATGAGCTGATTTCTCTTCCACTTGTATGCC 660
QY 858 AATGCTGTTATGAGAAATGTTCTTCTTCAATTAAGAGCCCTCAGAACCTTTGTA 917
Db 661 AATGCTGTTATGAGAAATGTTCTTCTTCAATTAAGAGCCCTCAGAACCTTTGTA 720
QY 918 CCAAGGCTGATGATCTTCCATTCGAGAGAACCTTTGATTTGATGAGATCACTTCAAG 977
Db 721 CCAAGGCTGATGATCTTCCATTCGAGAGAACCTTTGATTTGATGAGATCACTTCAAG 780
QY 978 AGTGTGACAGCAAAATGAAATCTGTGTTGTTATTTCTTGTGGGGCTGTATCAAG 1037
Db 781 AGTGTGACAGCAAAATGAAATCTGTGTTGTTATTTCTTGTGGGGCTGTATCAAG 840
QY 1038 CCAACATGATATGATGATTAACACATCTGGGAAACAAATGATGATGATCTTTCGG 1095
Db 841 CCAACATGATATGATGATTAACACATCTGGGAAATGATGATGATGATCTTTCGAG 898

RESULT 15
LOCUS CK320178 906 bp mRNA linear EST 11-MAY-2004

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 16:00:16 ; Search time 968.454 Seconds
(without alignments)
10758.117 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760

Sequence: 1 cctctcttatcctcttgc.....aatgtagtataatttgcgt 1760

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

N_Geneseq_16dec04:*
1: geneseqn19808:*
2: geneseqn19905:*
3: geneseqn20008:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1760	100.0	1760	12	ADQ14490 Wild type
2	1760	100.0	1782	2	AAV62440 Soybean w
3	1533	87.1	1533	12	ADQ14498 Mutant so
4	1533	87.1	1533	13	ADQ14498 Soybean m
5	1533	87.1	1533	13	ADQ14498 Soybean m
6	1531.4	87.0	1533	2	AAV62443 Soybean m
7	1531.4	87.0	1533	12	ADQ14494 Mutant so
8	1531.4	87.0	1533	13	ADQ14494 Soybean m
9	1465.8	83.3	1533	12	ADQ14504 Wild type
10	1465.8	83.3	1533	12	ADQ14504 Soybean m
11	1465.8	83.3	1533	13	ADQ14504 Soybean m
12	1465.8	83.3	1533	13	ADQ14504 Soybean m
13	1464.2	83.2	1533	13	ADQ14500 Mutant so
14	1464.2	83.2	1533	13	ADQ14500 Soybean m
15	1120.4	63.7	1950	2	AAV90402 Nicotiana
16	1001.4	56.9	1781	4	AAV87643 Brassica
17	997.6	56.7	1938	3	AAV50242 Arabidops
18	995.4	56.6	1533	6	AB213633 Arabidops
19	995.4	56.6	1533	12	ADN73524 Thale cre
20	994.6	56.5	1837	3	AAV34806 Arabidops

21	994.2	56.5	1959	4	AAV85922 MIP synth
22	990.8	56.3	1665	2	AAV09006 Inducible
23	990.4	56.3	1536	3	AAV34428 Arabidops
24	984	55.9	1931	2	AAV24407 Maize myo
25	956.4	54.3	1759	6	ABQ72653 Human MDD
26	835.8	47.5	1536	12	ADP43918 P. coarct
27	789.4	44.9	1719	3	AAV51095 Arabidops
28	552	31.4	1605	8	ABT19345 Aspergill
29	552	31.4	1605	8	ABT21165 Aspergill
30	514.4	29.2	2280	4	ABU17361 Drosophill
31	502.2	28.5	1772	4	AAH14171 Human CDN
32	502.2	28.5	1818	6	AAH15251 Human CDN
33	502.2	28.5	1825	6	ABT07176 Human ova
34	502.2	28.5	1852	4	AAH47741 Human hmt
35	502.2	28.5	1862	4	AAH59232 Human pol
36	502.2	28.5	2380	13	ADQ86789 Human tum
37	502.2	28.5	2380	13	ADQ85720 Human tum
38	502.2	28.5	2380	13	ACN40520 Tumour-as
39	487.2	27.7	1833	4	AAH47740 Murine nm
40	485	27.6	1991	12	ADG73735 Aspergill
41	472.4	26.8	618	13	ACN62890 Cotton ca
42	456.8	26.0	1835	12	ADQ25429 Human bof
43	454.4	25.8	1702	10	ADQ62765 Human CDN
44	451.2	25.6	2338	4	AAH161018 Human pol
45	441	25.1	1704	8	ABT20567 Aspergill

ALIGNMENTS

RESULT 1	
ID ADQ14490	
ID ADQ14490 standard; cDNA; 1760 BP.	
XX	
AC ADQ14490;	
XX	
DT 23-SEP-2004 (first entry)	
XX	
DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #1.	
XX	
KW Soybean, myo-inositol 1-phosphate synthase; gene; ss;	
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;	
KW inorganic phosphate.	
XX	
OS Glycine max.	
XX	
FH	
FT Key	Location/Qualifiers
FT CDS	54..1586
FT	/*tag= a
FT	/product= "wild type soybean myo-inositol 1-phosphate
FT	synthase #1"
XX	
PN US2004128713-A1.	
XX	
PD 01-JUL-2004.	
XX	
PF 21-NOV-2003; 2003US-00718952.	
XX	
PR 08-APR-1997; 97US-00835751.	
PR 07-APR-1998; 98WO-US006822.	
PR 26-APR-1999; 99US-00299315.	
PR 11-MAR-2002; 2002US-00025003.	
XX	
PA (HITZ/) HITZ W D.	
PA (SEBA/) SEBASTIAN S A.	
PA (GRAC/) GRACE D J.	
PA (STRE/) STREIT L G.	
XX	
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
XX	
DR WPI; 2004-533135/51.	
DR P-PSDB; ADQ14491.	
XX	

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX Claim 4; SEQ ID NO 1; 48pp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
CC
XX

Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Query Match 100.0%; Score 1760; DB 12; Length 1760;

Best Local Similarity 100.0%; Fred. No. 0;

Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTCTTATTCCTTTGTAATTCATTCATCTTAACTTTGTAATAATATGTTCA 60
DB 1 CTTCTCTTATTCCTTTGTAATTCATTCATCTTAACTTTGTAATAATATGTTCA 60
QY 61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTAGTACCGAGACTGATTCAGTCCG 120
DB 61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTAGTACCGAGACTGATTCAGTCCG 120
QY 121 TGTACACTAGCAAAACCGAATCTTTCACGAGAAAGAAATGGCACTATCAGTGA 180
DB 121 TGTACACTAGCAAAACCGAATCTTTCACGAGAAAGAAATGGCACTATCAGTGA 180
QY 181 TTGTCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTTAATTAG 240
DB 181 TTGTCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTTAATTAG 240
QY 241 GGGTAATGCTTGTGGGTTGGGTTGGAACACGCTCAACCTCAACCGGTGCTGTTATG 300
DB 241 GGGTAATGCTTGTGGGTTGGGTTGGAACACGCTCAACCTCAACCGGTGCTGTTATG 300
QY 301 CTAAACGAGAGGCAATTTATGGGCTAACAAAGACAAATTTCAAGGCAATTAATTG 360
DB 301 CTAAACGAGAGGCAATTTATGGGCTAACAAAGACAAATTTCAAGGCAATTAATTG 360
QY 361 GGTCCCTCAACCAAGCTTACGATTCGAGTTGGGCTTTCCAGGAGAGAAATCTATG 420
DB 361 GGTCCCTCAACCAAGCTTACGATTCGAGTTGGGCTTTCCAGGAGAGAAATCTATG 420
QY 421 CCCCATTTCAAGAGCTGCTTCCATAGTTAACCTTGACATGTTGTTGGGGATGGG 480
DB 421 CCCCATTTCAAGAGCTGCTTCCATAGTTAACCTTGACATGTTGTTGGGGATGGG 480
QY 481 AATATGACAACTGAACTGTGCTGATGCAATGGCCAGGCAAAAGTGTGATCAGATT 540
DB 481 AATATGACAACTGAACTGTGCTGATGCAATGGCCAGGCAAAAGTGTGATCAGATT 540
QY 541 TGCAGAGAGATTTGAGGCTTATCATGGAATCCATGTTCCATCCCGGAACTATGACC 600
DB 541 TGCAGAGAGATTTGAGGCTTATCATGGAATCCATGTTCCATCCCGGAACTATGACC 600
QY 601 CGGATTTCAATGCTGCAACAAAGAGAGCTGTCACCAACGTCATCAAGGGCACAAAGC 660

DB 601 CGGATTTCAATGCTGCAACAAAGAGAGCTGTCACCAACGTCATCAAGGGCACAAAGC 660
QY 661 AAGAGCAAGTTCAACCAATTCATCAAGACATCAAGGCGTTTAAAGAACCCAAAGTGG 720
DB 661 AAGAGCAAGTTCAACCAATTCATCAAGACATCAAGGCGTTTAAAGAACCCAAAGTGG 720
QY 721 ACAAGTGGTGTACTGTGGAATCTCTTGGCTGTGGAACAAGAAAGAGGCTGAGATTCTC 840
DB 721 ACAAGTGGTGTACTGTGGAATCTCTTGGCTGTGGAACAAGAAAGAGGCTGAGATTCTC 840
QY 781 TTAATGACCAATGAGAAATCTTGGCTGTGGAACAAGAAAGAGGCTGAGATTCTC 840
DB 781 TTAATGACCAATGAGAAATCTTGGCTGTGGAACAAGAAAGAGGCTGAGATTCTC 840
QY 841 CTTCACCTGTATGCAATGCTTGTGTATGAAAATGTTCTTCAATATGGAAGCC 900
DB 841 CTTCACCTGTATGCAATGCTTGTGTATGAAAATGTTCTTCAATATGGAAGCC 900
QY 901 CTCAAGACATTTTGTATCAAGGCTGATTTGATTCATGCGAGAACATTTGATTG 960
DB 901 CTCAAGACATTTTGTATCAAGGCTGATTTGATTCATGCGAGAACATTTGATTG 960
QY 961 GTGAGATGACTTCAAGAGTGTGTGACCAAAATGAATCTGTGTGTTGATTCTTGG 1020
DB 961 GTGAGATGACTTCAAGAGTGTGTGACCAAAATGAATCTGTGTGTTGATTCTTGG 1020
QY 1021 TGGGGGCTGATCAAGCCAACTATATGTCAGTTACATCAACATCTGGGAAACAATGATG 1080
DB 1021 TGGGGGCTGATCAAGCCAACTATATGTCAGTTACATCAACATCTGGGAAACAATGATG 1080
QY 1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGAAATCTCCAAAGACCAAGTGG 1140
DB 1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGAAATCTCCAAAGACCAAGTGG 1140
QY 1141 TTGATGATATGCTCAACGCAATGCTATGAGCTGTGTGAACATCAAGCATTG 1200
DB 1141 TTGATGATATGCTCAACGCAATGCTATGAGCTGTGTGAACATCAAGCATTG 1200
QY 1201 TTGATGATATGCTCAACGCAATGCTATGAGCTGTGTGAACATCAAGCATTG 1260
DB 1201 TTGATGATATGCTCAACGCAATGCTATGAGCTGTGTGAACATCAAGCATTG 1260
QY 1261 CTTCAAGATATTCATGAGTGGGAAAGAGACACCATTTGTTGCACAAACATGCGAGATT 1320
DB 1261 CTTCAAGATATTCATGAGTGGGAAAGAGAGACACCATTTGTTGCACAAACATGCGAGATT 1320
QY 1321 CCTCTTATGCTGCTCTTATTTATCTTGGACTGTGCTTCTTGTGAGCTCAGCACTAGAA 1380
DB 1321 CCTCTTATGCTGCTCTTATTTATCTTGGACTGTGCTTCTTGTGAGCTCAGCACTAGAA 1380
QY 1381 TCGAGTTTAAAGCTGAAAATGAGGAAATTCACATCTTCAACCCAGTTGCTACATCC 1440
DB 1381 TCGAGTTTAAAGCTGAAAATGAGGAAATTCACATCTTCAACCCAGTTGCTACATCC 1440
QY 1441 TCAGCTACCTGACCAAGGCTCTCTGTGTTCCACCGGTTACACAGTGTGATGCAATTGT 1500
DB 1441 TCAGCTACCTGACCAAGGCTCTCTGTGTTCCACCGGTTACACAGTGTGATGCAATTGT 1500
QY 1501 CAAGACGCTGCAATGCTGGAACCAATATGAGGCTGTGTGTGATGAGCCCAAGAGA 1560
DB 1501 CAAGACGCTGCAATGCTGGAACCAATATGAGGCTGTGTGTGATGAGCCCAAGAGA 1560
QY 1561 AATAATATCTCGAGATCAAGTGAACATGAGGACGAGAAATATATATGTTGGGGTATG 1620
DB 1561 AATAATATCTCGAGATCAAGTGAACATGAGGACGAGAAATATATATGTTGGGGTATG 1620
QY 1621 CCTAGCTGAATGTTTATGTTAATATATGTTGCTTAAATTTTGAAGTATATGAA 1680
DB 1621 CCTAGCTGAATGTTTATGTTAATATATGTTGCTTAAATTTTGAAGTATATGAA 1680
QY 1681 TGCATCACTTCATTAATGCTTTAGAGCGGAGCATATTCGTTTATGAGAAATGAAATG 1740

DB 1681 TGCATCAGCTTCAATATGCTTAGAGCGGCGCATATTCGTTTACTAGGAACATGAAATG 1740
QY 1741 AATGATGATTAATTTGTGT 1760
DB 1741 AATGATGATTAATTTGTGT 1760

RESULT 2
AAV62440
ID AAV62440 standard; cDNA; 1782 BP.
XX AAV62440;
AC 17-OCT-2003 (revised)
XX 02-FEB-1999 (first entry)
DT
XX Soybean wild-type myo-inositol 1-phosphate synthase cDNA.
DE
XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid; de.
XX
XX Glycine max; line LR13.
OS
XX
XX Key Location/Qualifiers
FH CDS 54..1586
FT /*tag= a
FT

XX W09845448-A1.
XX
XX 15-OCT-1998.
XX
XX 07-APR-1998; 98WO-US006822.
XX
XX 08-APR-1997; 97US-00835751.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 1998-568353/48.
XX P-PSDB; AAW79740.
XX
XX Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.
XX
XX
XX Example 5; Page 44-45; 63p; English.
XX
XX This is the nucleotide sequence of cDNA encoding the wild-type soybean
CC myo-inositol 1-phosphate synthase (MI 1-PS) present in clone pSdm1-1ps
CC (ATCC 97970). The clone was isolated from a cDNA library of soybean line
CC LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
CC thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,
CC raffinose and stachyose. A mutant MI 1-PS nucleic acid (see AAV62443) has
CC been identified in soybean line LR13, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC (G to T at base 1241) in the LR13 sequence. The mutation results in a
CC seed phenotype of very low raffinose saccharide sugars, very high sucrose
CC and low phytic acid. The nucleic acid is used to alter the raffinose
CC saccharide, sucrose, phytic acid and inorganic phosphate content of
CC soybean seeds, leading to useful soybean products, e.g. a seed phytic
CC acid content of less than 17 ug/g, a seed content of raffinose and
CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
CC
XX
XX Sequence 1782 BP; 516 A; 371 C; 411 G; 484 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 1760; DB 2; Length 1782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTTCTTATTCCTTTGTAATTCATTCATTCCTTAATCTTTGTAATAATTAATGTCA 60
|||||

DB 1 CTCTTCTTATTCCTTTGTAATTCATTCATTCCTTAATCTTTGTAATAATTAATGTCA 60
QY 61 TCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGACTGATTCAGTCCG 120
|||||
DB 61 TCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGACTGATTCAGTCCG 120
|||||

QY 121 TGTACAACTACGAAACCAACGAACTTGTACAGAGAACGGAATGACACTAATCAATGGA 180
121 TGTACAACTACGAAACCAACGAACTTGTGTACAGAGAACGGAATGACACTAATCAATGGA 180
DB 121 TGTACAACTACGAAACCAACGAACTTGTGTACAGAGAACGGAATGACACTAATCAATGGA 180
181 TTGTCAAAACCAATCTGTCAATAATGAAATTTAAACCAACATCCATGTTCTTAATTAG 240
DB 181 TTGTCAAAACCAATCTGTCAATAATGAAATTTAAACCAACATCCATGTTCTTAATTAG 240
QY 241 GGGTAAATGCTTGGGTTGGGTTGGGAACAGGCTCAACCTCACCGGTGTTATTG 300
241 GGGTAAATGCTTGGGTTGGGTTGGGAACAGGCTCAACCTCACCGGTGTTATTG 300
DB 241 GGGTAAATGCTTGGGTTGGGTTGGGAACAGGCTCAACCTCACCGGTGTTATTG 300
301 CTAAACGAGAGGCAATTCATGGGGCTACAAAGGACAAAGATTCAACAAAGCAATTAATTG 360
DB 301 CTAAACGAGAGGCAATTCATGGGGCTACAAAGGACAAAGATTCAACAAAGCAATTAATTG 360
QY 361 GCTCCCTCAACCAAGCCTCAGCTATCCAGTTGGGTCTTCCAGAGGAGAGAAATCTATG 420
361 GCTCCCTCAACCAAGCCTCAGCTATCCAGTTGGGTCTTCCAGAGGAGAGAAATCTATG 420
DB 421 CCCCATTCAGAGCCTGCTTCCAAATGTTAACCTCGACGACATGTTGTTGGGGATGGG 480
421 CCCCATTCAGAGCCTGCTTCCAAATGTTAACCTCGACGACATGTTGTTGGGGATGGG 480
DB 421 CCCCATTCAGAGCCTGCTTCCAAATGTTAACCTCGACGACATGTTGTTGGGGATGGG 480
481 AATATCAGCAACATGAACTGGGTGATGCCATGCGCAAGGGCAAAAGGTTTGACATGATT 540
DB 481 AATATCAGCAACATGAACTGGGTGATGCCATGCGCAAGGGCAAAAGGTTTGACATGATT 540
QY 541 TGCAGAGAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATCTATGACC 600
541 TGCAGAGAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATCTATGACC 600
DB 541 TGCAGAGAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATCTATGACC 600
601 CGGATTTCAATGCTGCGCAACCAAGAGAGCTGCCAACAGCTCATCAAGGCAACAAAGC 660
QY 601 CGGATTTCAATGCTGCGCAACCAAGAGAGCTGCCAACAGCTCATCAAGGCAACAAAGC 660
DB 601 CGGATTTCAATGCTGCGCAACCAAGAGAGCTGCCAACAGCTCATCAAGGCAACAAAGC 660
661 AAGAGCAAGTTTAAACAAATCATCAAAAGCATCAAGCCGTTTAAAGAACCAACAAAGTGG 720
QY 661 AAGAGCAAGTTTAAACAAATCATCAAAAGCATCAAGCCGTTTAAAGAACCAACAAAGTGG 720
DB 661 AAGAGCAAGTTTAAACAAATCATCAAAAGCATCAAGCCGTTTAAAGAACCAACAAAGTGG 720
721 ACAAGGTGTTGTAAGTGTGCACTGCAACACAGAGAGTACATTAATTTGTTGGGCGC 780
DB 721 ACAAGGTGTTGTAAGTGTGCACTGCAACACAGAGAGTACATTAATTTGTTGGGCGC 780
QY 781 TTAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAGATTCTC 840
781 TTAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAGATTCTC 840
DB 781 TTAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAGATTCTC 840
841 CTTTCACCTTGAATGCCATGCTTGTGTGTAAGAAATTTCTTTCAATTAATGAAAGCC 900
QY 841 CTTTCACCTTGAATGCCATGCTTGTGTGTAAGAAATTTCTTTCAATTAATGAAAGCC 900
DB 841 CTTTCACCTTGAATGCCATGCTTGTGTGTAAGAAATTTCTTTCAATTAATGAAAGCC 900
901 CTCAGAACACTTTTGTACAGAGGCTGATGATCTTGCCATCCGAGAGAACACTTGTATTG 960
QY 901 CTCAGAACACTTTTGTACAGAGGCTGATGATCTTGCCATCCGAGAGAACACTTGTATTG 960
DB 901 CTCAGAACACTTTTGTACAGAGGCTGATGATCTTGCCATCCGAGAGAACACTTGTATTG 960
961 GTGAGATGACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGTTGATTCTTGG 1020
QY 961 GTGAGATGACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGTTGATTCTTGG 1020
DB 961 GTGAGATGACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGTTGATTCTTGG 1020
1021 TGGGGGCTGATTCAGAGCAACATCTATATGATTAACAACATCTGGAACCAATGATG 1080
QY 1021 TGGGGGCTGATTCAGAGCAACATCTATATGATTAACAACATCTGGAACCAATGATG 1080
DB 1021 TGGGGGCTGATTCAGAGCAACATCTATATGATTAACAACATCTGGAACCAATGATG 1080
1081 GTATGAATCTTCCGAGCTCAACAACTTCCGTTCCAGAGAAATCTCCAGAGCAAGCTTG 1140
QY 1081 GTATGAATCTTCCGAGCTCAACAACTTCCGTTCCAGAGAAATCTCCAGAGCAAGCTTG 1140
DB 1081 GTATGAATCTTCCGAGCTCAACAACTTCCGTTCCAGAGAAATCTCCAGAGCAAGCTTG 1140

```
OY 1141 TTGATGATATGTCACAGCAATGCATCTCTATAGAGCCGTGTAACATCCAGCATG 1200
DB 1141 TTGAATGATATGTCACAGCAATGCATCTCTATAGAGCCGTGTAACATCCAGCATG 1200
OY 1201 TTGTTGTTATTAAGTATGTCCTTACGTAGGAGGACGAAAGAGCCATGATGATACA 1260
DB 1201 TTGTTGTTATTAAGTATGTCCTTACGTAGGAGGACGAAAGAGCCATGATGATACA 1260
OY 1261 CTTTCAGATATTCATGAGGTGGAAAGACCACTTTGTTGACAAACATCCGAGAT 1320
DB 1261 CTTTCAGATATTCATGAGGTGGAAAGACCACTTTGTTGACAAACATCCGAGAT 1320
OY 1321 CCCTCTAGCTGCTCCTATTTATCTTGAAGTTGATCTCTTCTAGAGCTCAGACATAGA 1380
DB 1321 CCCTCTAGCTGCTCCTATTTATCTTGAAGTTGATCTCTTCTAGAGCTCAGACATAGA 1380
OY 1381 TCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATTCATCCAGTTGCTACCATCC 1440
DB 1381 TCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATTCATCCAGTTGCTACCATCC 1440
OY 1441 TTAGCTACTCCACCAAGGTCCTCTGTGTTCCACCGGGTACACAGTGTGAATGCAATGT 1500
DB 1441 TTAGCTACTCCACCAAGGTCCTCTGTGTTCCACCGGGTACACAGTGTGAATGCAATGT 1500
OY 1501 CAAAGAGGCTGCAATGCTGAAAAACATATATGAGGCTTGTGATGAGCCCGCAGAGA 1560
DB 1501 CAAAGAGGCTGCAATGCTGAAAAACATATATGAGGCTTGTGATGAGCCCGCAGAGA 1560
OY 1561 ATACATGATTCCTGAGTACCAAGTGAAGATGGAACCAAGAAATATATATGTTGGGTAG 1620
DB 1561 ATACATGATTCCTGAGTACCAAGTGAAGATGGAACCAAGAAATATATATGTTGGGTAG 1620
OY 1621 CCTAGCTGAATGTTTATATATATATATATGTTTGTCTTAAATTTTGCAGTGAATGGA 1680
DB 1621 CCTAGCTGAATGTTTATATATATATATATGTTTGTCTTAAATTTTGCAGTGAATGGA 1680
OY 1681 TGCATCAGCTTCATTAATCTTTAGACGGGCAATTCCTGTTTACTAGAACATGAATG 1740
DB 1681 TGCATCAGCTTCATTAATCTTTAGACGGGCAATTCCTGTTTACTAGAACATGAATG 1740
OY 1741 AATGTAGTAAATTTTGTGT 1760
DB 1741 AATGTAGTAAATTTTGTGT 1760
```

```
RESULT 3
AD014498
ID AD014498 standard; cDNA; 1533 BP.
```

```
AD014498;
```

```
23-SEP-2004 (first entry)
```

```
Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.
```

```
KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KM inorganic phosphate; mutant.
```

```
OS Glycine max.
OS Synthetic.
```

```
XX Key Location/Qualifiers
```

```
FT CDS 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #2"
```

```
PN US2004128713-A1.
```

```
PD 01-JUL-2004.
```

```
XX
```

```
PF 21-NOV-2003; 2003US-00718952.
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI: 2004-53315/51.
XX P-PSDB; AD014499.
DR
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
PS Example 8; SEQ ID NO 9; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
CC
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
```

```
Query Match 87.1%; Score 1533; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 54 ATGTTTCATCAGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGATGAGATT 113
DB 1 ATGTTTCATCAGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGATGAGATT 60
OY 114 CAGTCCGTGTACACTTACGAAACCAACCGAATCTGTTTCACGAGAAACAGAAATGGCACTAT 173
DB 61 CAGTCCGTGTACACTTACGAAACCAACCGAATCTGTTTCACGAGAAACAGAAATGGCACTAT 120
OY 174 CAGTGAATTTGCAAAACCAAAATCTGTCAATTCGAATTTTAAACCAACATCCATGTTCT 233
DB 121 CAGTGAATTTGCAAAACCAAAATCTGTCAATTCGAATTTTAAACCAACATCCATGTTCT 180
OY 234 AAATTAGGGGTAAATGCTTTGAGGTTGGGGTGGAAACAACGGCTCAACCTCAACGGTGGT 293
DB 181 AAATTAGGGGTAAATGCTTTGAGGTTGGGGTGGAAACAACGGCTCAACCTCAACGGTGGT 240
OY 294 GTTATTTGCTAACCGAGAGGGCAATTCATGAGGCTACAAAGAGACAAATTCACACCAAT 353
DB 241 GTTATTTGCTAACCGAGAGGGCAATTCATGAGGCTACAAAGAGACAAATTCACACCAAT 300
OY 354 TACTTTGGTCCCTCAACCAAGCTCAGCTATCCAGATTGGGTCTTCCAGGAGAGAGAA 413
DB 301 TACTTTGGTCCCTCAACCAAGCTCAGCTATCCAGATTGGGTCTTCCAGGAGAGAGAA 360
```

```

Oy 414 ATCTATGCCCATTCAGAGCGCTTCCAAATGGTTAAACCTGACGACATGTGTTGGG 473
Db 361 ATTTATGCCCATTCAGAGCGCTTCCAAATGGTTAAACCTGACGACATGTGTTGGG 420
Oy 474 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGAGCAAGGTGTTGAC 533
Db 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGAGCAAGGTGTTGAC 480
Oy 534 ATCGATTTTGCAGAGCATTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 593
Db 481 ATCGATTTTGCAGAGCATTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 540
Oy 594 TATGACCCGGATTTCAATGCTGCCAACAGAGAGGTGCCAACACGTCATCAAGGCG 653
Db 541 TATGACCCGGATTTCAATGCTGCCAACAGAGAGGTGCCAACACGTCATCAAGGCG 600
Oy 654 ACAAGCAAGAGCATTTCAACAAATCATCAAGACATCAAGGCGTTTAAAGAAAGCAC 713
Db 601 ACAAGCAAGAGCATTTCAACAAATCATCAAGACATCAAGGCGTTTAAAGAAAGCAC 660
Oy 714 AAAGTGGACAAGGTGTTGTAAGTGGACTGCCAACAGAGGTACAGTAATTTGGTT 773
Db 661 AAAGTGGACAAGGTGTTGTAAGTGGACTGCCAACAGAGGTACAGTAATTTGGTT 720
Oy 774 GTGGGCGCTTAATGACCAATGGAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
Db 721 GTGGGCGCTTAATGACCAATGGAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
Oy 834 ATTTCTCTCTCCACCTTGTATGCGCATTTGTTGTTATGGAATAATTTCTTTCATTAAT 893
Db 781 ATTTCTCTCTCCACCTTGTATGCGCATTTGTTGTTATGGAATAATTTCTTTCATTAAT 840
Oy 894 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATTCCTTCAGAGAACACT 953
Db 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATTCCTTCAGAGAACACT 900
Oy 954 TTGATTTGGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTTGGTTGAT 1013
Db 901 TTGATTTGGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTTGGTTGAT 960
Oy 1014 TTCTTGTGGGGCTGTATCAAGCCCAATCTATAGTCATTAACAACACTGCGGAAAC 1073
Db 961 TTCTTGTGGGGCTGTATCAAGCCCAATCTATAGTCATTAACAACACTGCGGAAAC 1020
Oy 1074 AATGATGTATGAATCTTTGCGCTCCACAAACTTTCGTTCCAGGAAATCTTCCAGAGC 1133
Db 1021 AATGATGTATGAATCTTTGCGCTCCACAAACTTTCGTTCCAGGAAATCTTCCAGAGC 1080
Oy 1134 AACGTTGTGATGATATGTGTCACAGCAATGCAATCTCTATGAGCTGTGAACTCCA 1193
Db 1081 AACGTTGTGATGATATGTGTCACAGCAATGCAATCTCTATGAGCTGTGAACTCCA 1140
Oy 1194 GACCATGTTGTTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGAGCATGAT 1253
Db 1141 GACCATGTTGTTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGAGCATGAT 1200
Oy 1254 GAGTACACTTCAGAGATTTATCATGGGTGAAAAGACACATGTTTTTGCACACATGTC 1313
Db 1201 GAGTACACTTCAGAGATTTATCATGGGTGAAAAGACACATGTTTTTGCACACATGTC 1260
Oy 1314 GAGGATTCCTGTAGCTGCTCTATTAATCTTGGATTTGCTGCTTCTTGTGAGCTCAGC 1373
Db 1261 GAGGATTCCTGTAGCTGCTCTATTAATCTTGGATTTGCTGCTTCTTGTGAGCTCAGC 1320
Oy 1374 ACTAGATCGAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACACGATTTGCT 1433
Db 1321 ACTAGATCGAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACACGATTTGCT 1380
Oy 1434 ACCATCTCAGCTACTCTACCAAGGCTCTCTGTGTTCCACCGGGTACACACAGTGTGAAT 1493
Db 1381 ACCATCTCAGCTACTCTACCAAGGCTCTCTGTGTTCCACCGGGTACACACAGTGTGAAT 1440
Oy 1494 GCATTGTCAAACAGCGTGCATGCTGGAATAATAGGGCTTGTGATTTGGCC 1553

```

```

Db 1441 GCATTGTCAAACAGCGTGCATGCTGGAATAATAGGGCTTGTGATTTGGCC 1500
Oy 1554 CCAGAGATTAACATGATTTCTCGAGTACAGTGA 1586
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACAGTGA 1533

RESULT 4
ADS81999
ID ADS81999 standard; cDNA, 1533 BP.
XX
XX ADS81999;
AC
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
DE
XX
XX Soybean plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; line 29004JP01.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag=a
FT /product="myo-inositol 1-phosphate synthase"
PN US2003074685-A1.
XX
XX
PD 17-APR-2003.
XX
XX
XX 11-MAR-2002; 2002US-00025003.
PF
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX
XX Hitz WD, Sebastian SA;
PI
XX
XX WPI; 2004-639957/62.
DR P-PSDB; ADS82000.
XX
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
PT
XX
XX Example 8; SEQ ID NO 9; 34bp; English.
PS
XX
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 14.5 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a
CC

```

CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

XX
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Query Match 87.1%; Score 1533; DB 13; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ATGTTTCATCGAATAATTTTAAAGTTGAGTGTCTTAATGTGAATACCCGAGCTGAGAT 113
DB 1 ATGTTTCATCGAATAATTTTAAAGTTGAGTGTCTTAATGTGAATACCCGAGCTGAGAT 60
QY 114 CAGTCCGTTGACAACTACGAAACCAACCGAATCTGTTACGAGAACGAAATGGCACTAT 173
DB 61 CAGTCCGTTGACAACTACGAAACCAACCGAATCTGTTACGAGAACGAAATGGCACTAT 120
QY 174 CAGTGGATTGTCAAAACCCAAATCTGTCAATAAGAAATTTAAACCAACATCCATGTTCT 233
DB 121 CAGTGGATTGTCAAAACCCAAATCTGTCAATAAGAAATTTAAACCAACATCCATGTTCT 180
QY 234 AAATTAAGGGTAATGTTGTGGGTGGGGTGGAAACACGGCTCAACCTCCACGGTGT 293
DB 181 AAATTAAGGGTAATGTTGTGGGTGGGGTGGAAACACGGCTCAACCTCCACGGTGT 240
QY 294 GTTATTGCTAACCGAGAGGGCAATTTCAATGGGCTACAAAGAACAGATTCAACAGGCAT 353
DB 241 GTTATTGCTAACCGAGAGGGCAATTTCAATGGGCTACAAAGAACAGATTCAACAGGCAT 300
QY 354 TACTTTGGCTCCCTCAACCCAGGCTCAGTATCCGAGTTGGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAGGCTCAGTATCCGAGTTGGGTCTTCCAGGAGAGAA 360
QY 414 ATCTATGCCCCCATTCGAAGACCTGCTTCCAAATGTTTAAACCTGACGACATTTGTTGGG 473
DB 361 ATCTATGCCCCCATTCGAAGACCTGCTTCCAAATGTTTAAACCTGACGACATTTGTTGGG 420
QY 474 GGATGGGATATGACCAACATGAACCTGGCTGATGTCATGGCCAGGCAAAAGGTTTGAC 533
DB 421 GGATGGGATATGACCAACATGAACCTGGCTGATGTCATGGCCAGGCAAAAGGTTTGAC 480
QY 534 ATCGATTTGACAGAACGATTGAGGCTTTACATGAATCCATGCTTCCACTCCCGGAATC 593
DB 481 ATCGATTTGACAGAACGATTGAGGCTTTACATGAATCCATGCTTCCACTCCCGGAATC 540
QY 594 TATGACCCCGGATTTCTTCTGTCCCAACGAGAGACCGTCCAAACCTCATCAAGGGC 653
DB 541 TATGACCCCGGATTTCTTCTGTCCCAACGAGAGACCGTCCAAACCTCATCAAGGGC 600
QY 654 ACAAAAGCAGAGCAATTCACAAATCATCAAGACATCAAGGCGTTTAAAGAGCCACC 713
DB 601 ACAAAAGCAGAGCAATTCACAAATCATCAAGACATCAAGGCGTTTAAAGAGCCACC 660
QY 714 AAAAGTGACAGAGTGTGTTTACTGTGACTGCGCAACAGAGAGTACAGTAATTTGGTT 773
DB 661 AAAAGTGACAGAGTGTGTTTACTGTGACTGCGCAACAGAGAGTACAGTAATTTGGTT 720
QY 774 GTGGGCTTAATGACCAATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
DB 721 GTGGGCTTAATGACCAATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY 834 ATTTCTCTTCCACTTGTATGCACTGCTGTGTTATGGAATAATGTTCTTTCAATTAAT 893
DB 781 ATTTCTCTTCCACTTGTATGCACTGCTGTGTTATGGAATAATGTTCTTTCAATTAAT 840
QY 894 GGAAGCCCTCAGAACCTTTTGTACGAGGCTGATTTGATCTTGCATCGCAGGAACACT 953
DB 894 GGAAGCCCTCAGAACCTTTTGTACGAGGCTGATTTGATCTTGCATCGCAGGAACACT 953

DB 841 GGAAGCCCTCAGAACACTTTTGTACGAGGCTGATTTGATCTTGCATCGCAGGAACACT 900
QY 954 TTGATTTGGTGAAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTGTGGTTGAT 1013
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTGTGGTTGAT 960
QY 1014 TTCTTTGGGGGGCTGTATCAAGCCAACTCTATAGCAATTACACCATCTGGGAAAC 1073
DB 961 TTCTTTGGGGGGCTGTATCAAGCCAACTCTATAGCAATTACACCATCTGGGAAAC 1020
QY 1074 AATGATGATGAATCTTTGGGCTTCCAAACCTTCCGTTCCAGGAATCTCCAAAGAC 1133
DB 1021 AATGATGATGAATCTTTGGGCTTCCAAACCTTCCGTTCCAGGAATCTCCAAAGAC 1080
QY 1134 AACGTTGTGATGATAGTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCA 1193
DB 1081 AACGTTGTGATGATAGTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCA 1140
QY 1194 GACCATGTTGTTTATTAATGATGTGCTTACGTAAGGGAACAGCAAGAGCCATGGAT 1253
DB 1141 GACCATGTTGTTTATTAATGATGTGCTTACGTAAGGGAACAGCAAGAGCCATGGAT 1200
QY 1254 GAGTACACTTCAAGATATTTCAATGGGTGAAAGAGCAACATTTGTTGCAACAACATGC 1313
DB 1201 GAGTACACTTCAAGATATTTCAATGGGTGAAAGAGCAACATTTGTTGCAACAACATGC 1260
QY 1314 GAGATTCCTCTTATGCTGCTCTTATTAATCTTGAATGATGCTTCTTGTGAGCTCAGC 1373
DB 1261 GAGATTCCTCTTATGCTGCTCTTATTAATCTTGAATGATGCTTCTTGTGAGCTCAGC 1320
QY 1374 ACTAGAAATCGAATTTAAAGCTGAAATAGAGGAAATTCACATCTCCACCAGTGTCT 1433
DB 1321 ACTAGAAATCGAATTTAAAGCTGAAATAGAGGAAATTCACATCTCCACCAGTGTCT 1380
QY 1434 ACCATCTCAGTACTTCAACCAAGGCTCTCTGTTTCCACCGGGTACACCATGGTGAAT 1493
DB 1381 ACCATCTCAGTACTTCAACCAAGGCTCTCTGTTTCCACCGGGTACACCATGGTGAAT 1440
QY 1494 GCATTGTCAAAGCAGCTGCAATGCTGGAATAACATATAGAGGCTTGTGTGATTTGGCC 1553
DB 1441 GCATTGTCAAAGCAGCTGCAATGCTGGAATAACATATAGAGGCTTGTGTGATTTGGCC 1500
QY 1554 CCAGAGAAATTAACATGATTTCTGAGTACAAATGA 1586
DB 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAATGA 1533
RESULT 5
ADS81993
ID ADS81993 standard; cDNA, 1533 BP.
XX
AC ADS81993;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; starchose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag=a
FT /product="myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX


```
Db 1081 AACGTTGATGATATGTCACAGACANTGCCATCTCTATGAGCTGTGTAACATCCA 1140
Qy 1194 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGCCATGAT 1253
Db 1141 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAGGGGACAGCAATAGAGCCATGAT 1200
Qy 1254 GAGTACACTTCAGAGATATTCATGAGTGGGAAAAGAGACCAATGTTTGGACAAACATATC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGAGTGGGAAAAGAGACCAATGTTTGGACAAACATATC 1260
Qy 1314 GAGATTCCTCTGAGCTCTCTATTAATCTTGACATTTGCTCTTCTGAGCTACAG 1373
Db 1261 GAGGATTCCTCTGAGCTCTCTATTAATCTTGACATTTGCTCTTCTGAGCTACAG 1320
Qy 1374 ACTAGATGAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACACCAATGCT 1433
Db 1321 ACTAGATGAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACACCAATGCT 1380
Qy 1434 ACCATCTGAGTAACTCAACCAAGGCTCTGTTGTTCCACCGGTAACACAGTGTGAT 1493
Db 1381 ACCATCTGAGTAACTCAACCAAGGCTCTGTTGTTCCACCGGTAACACAGTGTGAT 1440
Qy 1494 GCATTGTCAAAGCAGGTCGAATGCTGAAAACATATGAGGCTTGTGATGGCC 1553
Db 1441 GCATTGTCAAAGCAGGTCGAATGCTGAAAACATATGAGGCTTGTGATGGCC 1500
Qy 1554 CCAGAGATTAACATGATTTCTGAGTACAAAGTGA 1586
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533
XX
```

RESULT 7

ADQ14494
ID ADQ14494 standard; cDNA; 1533 BP.

AC ADQ14494;

XX 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.

KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;

KM inorganic phosphate; raffinose; stachyose; phytic acid; sucrose;

OS inorganic phosphate; mutant.

OS Glycine max.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..1533

FT /tag= a

FT /product= "Mutant soybean myo-inositol 1-phosphate

FT synthase #1"

FT replace(1241,G)

FT mutation

FT /tag= b

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 26-APR-1999; 99US-0029315.

XX 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX (GRAC/) GRACE D J.

XX (STRE/) STREIT L G.

XX

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX MPI, 2004-533135/51.
DR F-PSDB; ADQ14495.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Claim 10; SEQ ID NO 5; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX

Seq Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 87.0%; Score 1531.4; DB 12; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 54 ATGTTTCATGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAATGATACCCGAGCTGAGAT 113
Db 1 ATGTTTCATGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAATGATACCCGAGCTGAGAT 60
Qy 114 CAGTCCGTTGTAACATCAACCAACCACTGTTTCAAGAAACAGAAATGGACCTAT 173
Db 61 CAGTCCGTTGTAACATCAACCAACCACTGTTTCAAGAAACAGAAATGGACCTAT 120
Qy 174 CAGTGGATTGCAAAACCAATCTGTCAATAGCAATTTAAACCAACATTCATGTTCT 233
Db 121 CAGTGGATTGCAAAACCAATCTGTCAATAGCAATTTAAACCAACATTCATGTTCT 180
Qy 234 AAATTAGGGGTATATGCTGTGGGTGGGGTGAACCAACGCTCAACCTGACCGGTGT 293
Db 181 AAATTAGGGGTATATGCTGTGGGTGGGGTGAACCAACGCTCAACCTGACCGGTGT 240
Qy 294 GTTATTTGCTAACCGAGAGGCAATTCAAGGCTCAACAAAGGACAAAGTTCAAAGCAAT 353
Db 241 GTTATTTGCTAACCGAGAGGCAATTCAAGGCTCAACAAAGGACAAAGTTCAAAGCAAT 300
Qy 354 TACTTTGGCTCCCTCAACCAACCTCAAGTATCCGATGGGTCTCTTCCAGGAGAGGAA 413
Db 301 TACTTTGGCTCCCTCAACCAACCTCAAGTATCCGATGGGTCTCTTCCAGGAGAGGAA 360
Qy 414 ATCTATGCCCACTTCAAGAGCTGCTTCCATAGTTTAAACCTTGACGACATTTGTGGG 473
Db 361 ATCTATGCCCACTTCAAGAGCTGCTTCCATAGTTTAAACCTTGACGACATTTGTGGG 420
Qy 474 GGATGGATATAGCAACATGAACCTGCTGATGTCATGCGCAGGAGCAAGGTGTTTAC 533
Db 421 GGATGGATATAGCAACATGAACCTGCTGATGTCATGCGCAGGAGCAAGGTGTTTAC 480
Qy 534 ATCGATTTGCAAGACAGTTGAGGCTTAAATGATGATCCATGCTTCACTCCCGGAATC 593
Db 481 ATCGATTTGCAAGACAGTTGAGGCTTAAATGATGATCCATGCTTCACTCCCGGAATC 540
```

```
OY 594 TATGACCCGGATTTTCATTGCTGCTCCACCAAGAGAGCGCTGCCAACACGTCATCAAGGCG 653
DB 541 TATGACCCGGATTTTCATTGCTGCTCCAAACCAAGAGAGCGCTGCCAACACGTCATCAAGGCG 600
OY 654 ACAAAGCAAGAGCAAGTTCACAAATTCATCAAGAGATCAAGCGCTTTAAGAACCCACC 713
DB 601 ACAAAGCAAGAGCAAGTTCACAAATTCATCAAGAGATCAAGCGCTTTAAGAACCCACC 660
OY 714 AAAGTGACAAAGGTGTTGACTGTGACCTGCAACACAGAGAGGTACAGTAATTTGGTT 773
DB 661 AAAGTGACAAAGGTGTTGACTGTGACCTGCAACACAGAGAGGTACAGTAATTTGGTT 720
OY 774 GTGGGCTTTAATGACACATGAGAAATCTCTTGCTGCTGTGACAGAAATAGGCTGAG 833
DB 721 GTGGGCTTTAATGACACATGAGAAATCTCTTGCTGCTGTGACAGAAATAGGCTGAG 780
OY 834 ATTTCTCTTCCACCTTGATGCCATGCTGTGTGTTAAGAAATGTTCTTCTTCAATAT 893
DB 781 ATTTCTCTTCCACCTTGATGCCATGCTGTGTGTTAAGAAATGTTCTTCTTCAATAT 840
OY 894 GGAGCCCTCAGAACACTTTGTACAGGGCTGATGATCTTGCCATGCGAGAACACT 953
DB 841 GGAGCCCTCAGAACACTTTGTACAGGGCTGATGATCTTGCCATGCGAGAACACT 900
OY 954 TTGATTTGTGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGAT 1013
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGAT 960
OY 1014 TTCTCTTGAGGGGCTGTGATCAAGCCAAATCTATATGTCAGTTAAACCATTTGGGAAAC 1073
DB 961 TTCTCTTGAGGGGCTGTGATCAAGCCAAATCTATATGTCAGTTAAACCATTTGGGAAAC 1020
OY 1074 AATGATGTGATGAATCTTTCGGCTCCACAACTTTCGGTCCAGAGAAATCTCAAGAGC 1133
DB 1021 AATGATGTGATGAATCTTTCGGCTCCACAACTTTCGGTCCAGAGAAATCTCAAGAGC 1080
OY 1134 AAGTTTGTGATGATATGTCACACGAATGCCATCTCTATGAGCTGTGTAACATCCA 1193
DB 1081 AAGTTTGTGATGATATGTCACACGAATGCCATCTCTATGAGCTGTGTAACATCCA 1140
OY 1194 GAGCAGTGTGTTATTAAGATGTCCTTACGTAAGGGGACAGACAGAGACCCATGAT 1253
DB 1141 GAGCAGTGTGTTATTAAGATGTCCTTACGTAAGGGGACAGACAGAGACCCATGAT 1200
OY 1254 GAGTACACTTCAAGATATTCATGAGGTGAGAAAGAGACCATTTGTTGCAACAACATGC 1313
DB 1201 GAGTACACTTCAAGATATTCATGAGGTGAGAAAGAGACCATTTGTTGCAACAACATGC 1260
OY 1314 GAGATTCCTCTTACGCTCTCTATATCTTGAAGCTGTCCTTCTGCTGAGCTCAGC 1373
DB 1261 GAGATTCCTCTTACGCTCTCTATATCTTGAAGCTGTCCTTCTGCTGAGCTCAGC 1320
OY 1374 ACTAAGATGAGATTAAGCTGAGAAATGAGGAAATTCACATCTCCATCCACAGTGGT 1433
DB 1321 ACTAAGATGAGATTAAGCTGAGAAATGAGGAAATTCACATCTCCATCCACAGTGGT 1380
OY 1434 ACCATCTCAGGTACCTCAACAGGCTCTCTGTTTCCACCGGATACCAAGTGTGAAT 1493
DB 1381 ACCATCTCAGGTACCTCAACAGGCTCTCTGTTTCCACCGGATACCAAGTGTGAAT 1440
OY 1494 GCATTGTCAAAAGCAGGTGCATATGCTGAAACATATATGAGGGCTTGTGTGATTGGCC 1553
DB 1441 GCATTGTCAAAAGCAGGTGCATATGCTGAAACATATATGAGGGCTTGTGTGATTGGCC 1500
OY 1554 CCAGAGAAATACATGATTTCTGAGATCAAGTGA 1586
DB 1501 CCAGAGAAATACATGATTTCTGAGATCAAGTGA 1533
```

RESULT 8
ADS81997
ID ADS81997 strand; cDNA; 1533 BP.

```
AC ADS81997;  
XX 18-NOV-2004 (first entry)  
DT  
XX  
DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.  
XX  
XX Soybean; plant; myo-inositol 1-phosphate synthase; ssf gene; phytic acid;  
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.  
XX  
XX Glycine max; line LR33.  
OS  
XX Synthetic.  
FH  
FH Key Location/Qualifiers  
FT 1..1533  
FT CDS /*tag=a  
FT /product="myo-inositol 1-phosphate synthase"  
FT replace(1188,G)  
FT /*tag=b  
FT mutation  
FT  
XX  
XX US2003074685-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 11-MAR-2002; 2002US-00025003.  
XX  
XX 08-APR-1997; 97US-00835751.  
XX 07-APR-1998; 98WO-US006822.  
XX  
XX (HITZ/) HITZ W D.  
XX (SEBA/) SEBASTIAN S A.  
XX  
XX Hitz WD, Sebastian SA;  
PI  
XX  
XX WPI: 2004-639957/62.  
XX P-PsDB; ADS81998.  
XX  
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-  
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,  
PT phytic acid and inorganic phosphate content of soybean seeds.  
PT  
XX  
XX Claim 8; SEQ ID NO 5; 34pp; English.  
XX  
XX The invention relates to an isolated nucleic acid fragment encoding a  
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-  
CC phosphate synthase having decreasing capacity for the synthesis for myo-  
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the  
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its  
CC complement, subfragment or the complement of the subfragment, operably  
CC linked to, suitable regulatory sequences, where expression of the chimeric  
CC gene results in a decrease in expression of an endogenous or native gene  
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant  
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic  
CC acid content of less than 17 micromol/g, a seed content of raffinose plus  
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of  
CC greater than 200 micromol/g), provided that the plant is not LR33), seeds  
CC from the plant, making a soybean plant with the heritable phenotype  
CC (comprising crossing LR33 or the plant comprising the chimeric gene with  
CC an elite soybean plant and selecting a progeny plant of the cross of  
CC crossing step that has a heritable phenotype as mentioned above), seeds  
CC of soybean plant made by the above method, a soy protein product derived  
CC from seeds of a soybean plant (homozygous for one or more gene encoding a  
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for  
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a  
CC heritable phenotype as mentioned above), and making or producing a  
CC soybean protein product derived from seeds of a soybean plant with  
CC heritable phenotype as mentioned above. The nucleic acid is useful for  
CC altering raffinose saccharide, sucrose, phytic acid and inorganic  
CC phosphate content of soybean seeds thus leading to valuable and useful  
CC soybean products, since the presence of high concentration of raffinose  
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence  
CC when consumed by humans. The present sequence encodes a mutant myo-  
XX inositol 1-phosphate synthase.
```

SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
 Query Match 87.0%; Score 1531.4; DB 13; Length 1533;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	54	ATGTCATCGAGAAATTTAAAGTTGAGTGTCTAATGTGAAGTACCGAGCTGAGATT	113
DB	1	ATGTCATCGAGAAATTTAAAGTTGAGTGTCTAATGTGAAGTACCGAGCTGAGATT	60
QY	114	CAGTCCGTGTACAACTACGAACACCGAATTGTTACGAGAACAGAAATGACCTAT	173
DB	61	CAGTCCGTGTACAACTACGAACACCGAATTGTTACGAGAACAGAAATGACCTAT	120
QY	174	CAGTGAATGTCAAAACCCAAATCTGTCAATAGAAATTTAAACCAATCCATGTTCT	233
DB	121	CAGTGAATGTCAAAACCCAAATCTGTCAATAGAAATTTAAACCAATCCATGTTCT	180
QY	234	AAATTAGGGGTAAATGCTTGTGGGTGGGTGAGAAACAAGGCTCAACCCCTCACGGTGT	293
DB	181	AAATTAGGGGTAAATGCTTGTGGGTGGGTGAGAAACAAGGCTCAACCCCTCACGGTGT	240
QY	294	GTTATTTGCTAACCGAGAGGGCATTTTCAATGAGCTACAAAGACAAGATTCAACAGCAAT	353
DB	241	GTTATTTGCTAACCGAGAGGGCATTTTCAATGAGCTACAAAGACAAGATTCAACAGCAAT	300
QY	354	TACTTTGGCTCCTTCAACCCAGCTCAGCTATCCGAGTTGGTCTTCCAGGGAGAGAA	413
DB	301	TACTTTGGCTCCTTCAACCCAGCTCAGCTATCCGAGTTGGTCTTCCAGGGAGAGAA	360
QY	414	ATCTATGCCCATTCAGAGGCTGCTTCCAAATGTTAACCTGACGACATGTTGTTGG	473
DB	361	ATCTATGCCCATTCAGAGGCTGCTTCCAAATGTTAACCTGACGACATGTTGTTGG	420
QY	474	GATGGGATATACGAACAATGAACCTGTGATGCCATGGCCAGGCAAAAGTGTGAC	533
DB	421	GATGGGATATATCGAACAATGAACCTGTGATGCCATGGCCAGGCAAAAGTGTGAC	480
QY	534	ATCGATTTGACAGACGATTGAGGCTTACATGGAATCCATGCTTCCACTCCCGGAATC	593
DB	481	ATCGATTTGACAGACGATTGAGGCTTACATGGAATCCATGCTTCCACTCCCGGAATC	540
QY	594	TATGACCCGGATTTCATGCTGCCCAACAGAGAGGTGCCAAACGTCATCAAGGGC	653
DB	541	TATGACCCGGATTTCATGCTGCCCAACAGAGAGGTGCCAAACGTCATCAAGGGC	600
QY	654	ACAAAGCAAGAGCAATTCACAAATCATCAAGACATCAAGGCTTTAAGAAAGCACCC	713
DB	601	ACAAAGCAAGAGCAATTCACAAATCATCAAGACATCAAGGCTTTAAGAAAGCACCC	660
QY	714	AAAGTGAACAGGTGTTGTAATCTGTGACTGCCAACAAGAGGTACAGTAATTTGGTT	773
DB	661	AAAGTGAACAGGTGTTGTAATCTGTGACTGCCAACAAGAGGTACAGTAATTTGGTT	720
QY	774	GTGGGCTTAAATGACCAATGAGAAATCTTGGCTGTGACAGAAATGAGGCTGAG	833
DB	721	GTGGGCTTAAATGACCAATGAGAAATCTTGGCTGTGACAGAAATGAGGCTGAG	780
QY	834	ATTTCTCTTCCACTTGTATGTCATGCTGTGTATGGAATAATGCTTCCATTAAAT	893
DB	781	ATTTCTCTTCCACTTGTATGTCATGCTGTGTATGGAATAATGCTTCCATTAAAT	840
QY	894	GGAAAGCCTCAGAACCTTTTGTACAGAGGCTGATTTGTCATCGGAGAACACT	953
DB	841	GGAAAGCCTCAGAACCTTTTGTACAGAGGCTGATTTGTCATCGGAGAACACT	900
QY	954	TTGATTTGGTGGAGATGATCTTCAAGAGTGTGACCAAAATGAATACTGTGTTGTTAT	1013
DB	901	TTGATTTGGTGGAGATGATCTTCAAGAGTGTGACCAAAATGAATACTGTGTTGTTAT	960
QY	1014	TTCTCTTGGGGGCTGTATCAAGCCAACTATATAGTCAAGTACCAACCATCTGGGAAAC	1073
DB	961	TTCTCTTGGGGGCTGTATCAAGCCAACTATATAGTCAAGTACCAACCATCTGGGAAAC	1020

QY	1074	AAATGATGATGAATCTTCCGGCTCCACAAACCTTCCGTTCCAAAGGAATCTCCAGAGC	1133
DB	1021	AAATGATGATGAATCTTCCGGCTCCACAAACCTTCCGTTCCAAAGGAATCTCCAGAGC	1080
QY	1134	AACGTTGTGATGATATGCTCAACAGCAATGCAATGCTCTATGAGGCTGTGAATCCA	1193
DB	1081	AACGTTGTGATGATATGCTCAACAGCAATGCAATGCTCTATGAGGCTGTGAATCCA	1140
QY	1194	GACCATGTTGTTATTAATGATGTGCTTACATGAGGAGACAGCAAGAGCCATGAT	1253
DB	1141	GACCATGTTGTTATTAATGATGTGCTTACATGAGGAGACAGCAATGAGCCATGAT	1200
QY	1254	GAGTACCTTCAGAGATATTCATGAGTGGAGAAAGACACATGTTTGAACAACATGC	1313
DB	1201	GAGTACCTTCAGAGATATTCATGAGTGGAGAAAGACACATGTTTGAACAACATGC	1260
QY	1314	GAGGATTTCCCTCTAGAGCTCTCTATTAATCTTGAACCTTGCTCTTGAGCTCAGC	1373
DB	1261	GAGGATTTCCCTCTAGAGCTCTCTATTAATCTTGAACCTTGCTCTTGAGCTCAGC	1320
QY	1374	ACTAGATCGAGTTTAAAGCTGAAAATGAGGAAAATTCACCTATTCACCCAGTTGCT	1433
DB	1321	ACTAGATCGAGTTTAAAGCTGAAAATGAGGAAAATTCACCTATTCACCCAGTTGCT	1380
QY	1434	ACCATCTCAGCTACTCAACAAAGGCTCTCTGTGTTCCACGGGTACACAGTGTGAAT	1493
DB	1381	ACCATCTCAGCTACTCAACAAAGGCTCTCTGTGTTCCACGGGTACACAGTGTGAAT	1440
QY	1494	GCATTGTCAAGCAGGCTGCAATGCTGAAAACATTAATGAGGCTTGTGTGATTGGCC	1553
DB	1441	GCATTGTCAAGCAGGCTGCAATGCTGAAAACATTAATGAGGCTTGTGTGATTGGCC	1500
QY	1554	CCAGAGATTAACATGATTCTCGAGTACAAAGTGA	1586
DB	1501	CCAGAGATTAACATGATTCTCGAGTACAAAGTGA	1533

RESULT 9
 ADQ14504
 ID ADQ14504 standard; cDNA, 1533 BP.
 XX
 AC ADQ14504;
 XX
 DT 23-SBP-2004 (first entry)
 XX
 DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.
 XX
 KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 XX inorganic phosphate.
 XX
 OS Glycine max.
 XX
 FH
 FT Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "Wild type soybean myo-inositol 1-phosphate
 FT synthase #2"
 XX
 PN US2004128713-A1.
 XX
 PD 01-JUL-2004.
 XX
 PD 21-NOV-2003; 2003US-00718952.
 XX
 PF 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98MO-US006822.
 PR 26-APR-1999; 99US-00293315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI: 2004-53335/51.
DR P-PSDB; ADQ14505.

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Claim 4; SEQ ID NO 15; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 83.3%; Score 1465.8; DB 12; Length 1533;

Best Local Similarity 97.3%; Pred. No. 0;

Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 54 ATGTCATGAGAAATTTTAAAGTTGAGTCTTAATGTAAGTACACGAGCTGAGATT 113
DB 1 ATGTCATGAGAAATTTTAAAGTTGAGTCTTAATGTAAGTACACGAGCTGAGATT 60
QY 114 CAGTCGCTGTAACAATGAGAAACCAAGATTGTTCAAGAAACGAATGAGCACTAT 173
DB 61 CAGTCGCTGTAACAATGAGAAACCAAGATTGTTCAAGAAACGAATGAGCACTAT 120
QY 174 CAGTGGATTGTCAAAACCAATCTGTCAATAAGAAATTTAAACCAACATCGATTCT 233
DB 121 CAGTGGATTGTCAAAACCAATCTGTCAATAAGAAATTTAAACCAACATCGATTCT 180
QY 234 AAATTTAAGGGTAATGTTGTGGTTGGGGTGGAAACAAGGCTCAACCTTCACCGTGT 293
DB 181 AAATTTAAGGGTAATGTTGTGGTTGGGGTGGAAACAAGGCTTCACCTTCACCGTGT 240
QY 294 GTTATTTGTAACAGAGGCAATTTTCAAGGCTTACAAAGCAAGATTTCAACAGCAAT 353
DB 241 GTTATTTGTAACAGAGGCAATTTTCAAGGCTTACAAAGCAAGATTTCAACAGCAAT 300
QY 354 TACTTTGGCTCCCTCAACCCAAAGCTCAGCTATCCGAGTTGGGCTTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAAAGCTCAGCTATCCGAGTTGGGCTTTCCAGGAGAGAA 360
QY 414 ATCTATGCCCCATTTCAAGGCTGCTTCCAATGTTTAAACCTGACACATTTGTGTTGGG 473
DB 361 ATCTATGCCCCATTTCAAGGCTGCTTCCAATGTTTAAACCTGACACATTTGTGTTGGG 420
QY 474 GGATGGATATGAGCAACATGAACCTGGCTGATGCGATGGCGGCAAGGAGGTTTTCAC 533
DB 421 GGATGGATATGAGCAACATGAACCTGGCTGATGCGATGGCGGCAAGGAGGTTTTCAC 480
QY 534 ATCGATTTCAGAAAGCAGTTGAGGCTTACATGAATCCATGCTTCCCTCCCGGAATC 593

DB 481 ATCGATTTCAGAAAGCAGTTGAGGCTTACATGAATCATGTTCCACTCCCGGAATC 540
QY 594 TATGACCCGGGAATTTGATTGCTGCCAAGCAAGGAGCGTGCAACATGTCATCAAGGAC 653
DB 541 TACGACCCGGGAATTTGATTGCTGCCAAGCAAGGAGCGTGCCACACATGTTTAAAGGAC 600
QY 654 ACAAGCAAGGCAAGTTCAACAAATCATCAAGACATCAAGCGCTTTAAGAAACCAAC 713
DB 601 ACAAGCAAGGCAAGTTCAAGAAATCATCAAGACATCAAGCGCTTTAAGAAACCAAC 660
QY 714 AAATGGAACAAGGTGTTGTAAGTGTGAGCTGCCAACAAGAGAGGTAAAGTAATTTGGTT 773
DB 661 AAATGGAACAAGGTGTTGTAAGTGTGAGCTGCCAACAAGAGAGGTAAAGTAATTTGGTT 720
QY 774 GTGGGCTTTAATGACACCATGAGAAATCTCTGGCTGTGAGCAAGAAATGAGGCTGAG 833
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTCTGGCTGTGAGCAAGAAATGAGGCTGAG 780
QY 834 ATTTCTCCTTCACCTTGTATGCAATTCCTTGTGTATGAAATGTTCTTTCAATTAAT 893
DB 781 ATTTCTCCTTCACCTTGTATGCAATTCCTTGTGTATGAAATGTTCTTTCAATTAAT 840
QY 894 GGAAGCCTTCAGAACACTTTTGTATCAGAGGCTGATTGATCTTGGCATGCGAGAACT 953
DB 841 GGAAGCCTTCAGAACACTTTTGTATCAGAGGCTGATTGATCTTGGCATGCGAGAACT 900
QY 954 TTGATTGTGAGATGATCTTCAAGAGTGTGACCAACAAATGAATCTGTGTGTTGAT 1013
DB 901 TTGATTGTGAGATGATCTTCAAGAGTGTGACCAACAAATGAATCTGTGTGTTGAT 960
QY 1014 TTCTCTTGAGGAGGCTGTATGACCAACATCTATAGTCAATTACCAACATCTGGGAAAC 1073
DB 961 TTCTCTTGAGGAGGCTGTATGACCAACATCTATAGTCAATTACCAACATCTGGGAAAC 1020
QY 1074 AATGATGATGATATCTTTGGCTTCACAACTTTCCGTTCCAGAAATCTCCAAAGAC 1133
DB 1021 AATGATGATGATATCTTTGGCTTCACAACTTTCCGTTCCAGAAATCTCCAAAGAC 1080
QY 1134 AACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
DB 1081 AACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1194 GACCATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
DB 1141 GACCATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1254 GAGTACACTTCAGATATTTCAATGAGTGAAGAAAGACATCTTTTGTGCAACAATGC 1313
DB 1201 GAGTACACTTCAGATATTTCAATGAGTGAAGAAAGACATCTTTTGTGCAACAATGC 1260
QY 1314 GAGATTCCTCTTGAAGCTGCTCTATTAATCTTGAAGCTTGTGCTTGTGAGTCAAGC 1373
DB 1261 GAGATTCCTCTTGAAGCTGCTCTATTAATCTTGAAGCTTGTGCTTGTGAGTCAAGC 1320
QY 1374 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTTCACTCAATTCACCCAGTTGCT 1433
DB 1321 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTTCACTCAATTCACCCAGTTGCT 1380
QY 1434 ACCATCTCAGTATCTTCAACAGGCTCTCTGTTTCCACGGGTATCACAGTGTGAT 1493
DB 1381 ACCATCTCAGTATCTTCAACAGGCTCTCTGTTTCCACGGGTATCACAGTGTGAT 1440
QY 1494 GCATTGTCAAAAGCAGGTCAGATGCTGAGAAACATTAAGAGGCTTGTGATTTGGCC 1553
DB 1441 GCATTGTCAAAAGCAGGTCAGATGCTGAGAAACATTAAGAGGCTTGTGATTTGGCC 1500
QY 1554 CCAGAGAAATTAACATGATTTCTGAGTACAGTGA 1586
DB 1501 CCAGAGAAATTAACATGATTTCTGAGTACAGTGA 1533

RESULT 10

ADQ14502
ID ADQ14502 standard; cDNA, 1533 BP.
XX
AC ADQ14502;
XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
XX
KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
OS Glycine max.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #4"
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PP 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1996; 98WO-US006822.
PR 26-APR-1999; 99US-0029315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG,
XX
DR WPI; 2004-533135/51.
DR P-PSDB; ADQ14503.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Example 8; SEQ ID NO 13; 48pp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the plant,
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 83.3%; Score 1465.8; DB 12; Length 1533;
Best Local Similarity 97.3%; Pred. No. 0;

	Matches 1491;	Conservative 0;	Mismatches 42;	Indels 0;	Gaps 0;
QY	54	ATGTTTATGAGAAATTTAAAGTTGAGTGTCTTAATGTGAATGACCCGAGACTGAGATT	113		
DB	1	ATGTTTATGAGAAATTTAAAGTTGAGTGTCTTAATGTGAATGACCCGAGACTGAGATT	60		
QY	114	CAGTCCGTTACAACTACGAAACCAACCGAACTTTGTTACGAGAACGGAATGACCTAT	173		
DB	61	CAGTCCGTTACAACTACGAAACCAACCGAACTTTGTTACGAGAACGGAATGACCTAT	120		
QY	174	CAGTGAATGTCAAAACCAATCTGTCAATATGCAATTTAAACCAACATTCATGTTCT	233		
DB	121	CAGTGAATGTCAAAACCAATCTGTCAATATGCAATTTAAACCAACATTCATGTTCTCA	180		
QY	234	AAATTAGGGGTAATGTTGTGGTTGGGTTGAAACACGCTCAACCTCACGGTGGT	293		
DB	181	AAATTAGGGGTAATGTTGTGGTTGGGTTGAAACACGCTCAACCTCACGGTGGT	240		
QY	294	GTTATTGCTAACCGAGGGGCAATTCATGGGCTAACAAAGACAAAGATTCAACAGCCAT	353		
DB	241	GTTATTGCTAACCGAGGGGCAATTCATGGGCTAACAAAGACAAAGATTCAACAGCCAT	300		
QY	354	TACTTGGCTCCCTCACCCAAAGCTCAGCTATCCGAGTTGGGCTTCCAGGGAGAGAA	413		
DB	301	TACTTGGCTCCCTCACCCAAAGCTCAGCTATTCGAGTTGGATCTTCCAGGGAGAGAA	360		
QY	414	ATCTATGCCCAATTCAAGCCTGCTTCCAATGTTTAAACCTGACGACATTTGTGTTGG	473		
DB	361	ATCTATGCCCAATTCAAGCCTGCTTCCAATGTTTAAACCTGACGACATTTGTGTTGG	420		
QY	474	GGATGGGATATGACCAACATGAACCTGCTGATGTCATGCGCAGGCAAAAGGTTTGAC	533		
DB	421	GGATGGGATATGACCAACATGAACCTGCTGATGTCATGCGCAGGCAAAAGGTTTGAC	480		
QY	534	ATCGAATTTGCAAGACAGTTGAGGCTTACATGSAATCATGCTTCCATCCCGGAATC	593		
DB	481	ATCGAATTTGCAAGACAGTTGAGGCTTACATGSAATCATGCTTCCATCCCGGAATC	540		
QY	594	TATGACCCGGAATTTATGCTGCTCCAAACCAAGAGAGCGTGCCAAACGTATCAAGGCG	653		
DB	541	TATGACCCGGAATTTATGCTGCTCCAAACCAAGAGAGCGTGCCAAACGTATTAAGGCG	600		
QY	654	ACAAAGCAAGACCAAGTTCAAAATCATCAAGACATCAAGCGTTTAAAGAACGACAC	713		
DB	601	ACAAAGCAAGACCAAGTTCAAGAAATCATCAAGACATCAAGCGTTTAAAGAACGACAC	660		
QY	714	AAAGTGAACAAGGTGTTGTAAGTGAAGTGAACACAGAGAGGATGTAATTTGGTT	773		
DB	661	AAAGTGAACAAGGTGTTGTAAGTGAAGTGAACACAGAGAGGATGTAATTTGGTT	720		
QY	774	GTTGGCTTTAATGACACATGAGAAATCTTTGGCTGTGAGACAGAAATGAGGCTGAG	833		
DB	721	GTTGGCTTTAATGACACATGAGAAATCTTTGGCTGTGAGACAGAAATGAGGCTGAG	780		
QY	834	ATTTTCTCTTCCACCTTGTATGCCATGTGTTGTATGAGAAATGTTCTTTCAATAT	893		
DB	781	ATTTTCTCTTCCACCTTGTATGCCATGTGTTGTATGAGAAATGTTCTTTCAATAT	840		
QY	894	GGAAGCCCTCAGAACACTTTTATACAGGCGTATGATCTTGCCATGCGAGAACACT	953		
DB	841	GGAAGCCCTCAGAACACTTTTATACAGGCGTATGATCTTGCCATGCGAGAACACT	900		
QY	954	TTGATTTGTGAGATGACTTCAAGAGTGTCAAGCAAAATGTAATCTGTGTTGTTGAT	1013		
DB	901	TTGATTTGTGAGATGACTTCAAGAGTGTCAAGCAAAATGTAATCTGTGTTGTTGAT	960		
QY	1014	TTCTTTTGGGGGCTGTATCAAGCAACTTTATATGCTAGTTAAACAATTTGGGAAAC	1073		
DB	961	TTCTTTTGGGGGCTGTATCAAGCAACTTTATATGCTAGTTAAACAATTTGGGAAAC	1020		
QY	1074	AATGATGATGATGAATCTTTGGGCTCCAAAACTTTCGTTCCAAAGAAATCTCAAGAGC	1133		
DB	1021	AATGATGATGATGAATCTTTGGGCTCCAAAACTTTCGTTCCAAAGAAATCTCAAGAGC	1080		


```
OY 1134 AACGTTGTGATGATGTCGAACAGCAATGCCATCCTCTATAGACCTGTGAACATGCA 1193
DB 1081 AACGTTGTGATGATGTCGAACAGCAATGCCATCCTCTATAGACCTGTGAACATGCC 1140
OY 1194 GACCAATGTTGTTGTTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGCCATGAT 1253
DB 1141 GACCAATGTTGTTGTTATTAAGTATGTCCTTACGTAGGGGATAGCAAGAGCCATGAT 1200
OY 1254 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACCATTTGTTTGACAAACATGC 1313
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATTTGTTTGACAAACATGT 1260
OY 1314 GAGGATCCCTCTGATGCTCTCTATTTATCTTGACCTTGCTCTTGAGCTGACG 1373
DB 1261 GAGGATCCCTCTGATGCTCTCTATTTATCTTGAGCTTGCTCTTGCTGACGCTGACG 1320
OY 1374 ACTAGATCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCCTCAATCCACCACTTGCT 1433
DB 1321 ACTAGATCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCCTCAATCCACCACTTGCT 1380
OY 1434 ACCATCTCTAGCTACCTCAACCAAGCTCTCTGCTTCCACCGGGTACACCACTGCTGAT 1493
DB 1381 ACCATCTCTAGCTACCTCAACCAAGCTCTCTGCTTCCACCGGGTACACCACTGCTGAT 1440
OY 1494 GCATTGTCAAAGCAGCGTGCATGCTGAAAAATCATATGAGGGCTTGTGGATTGGCC 1553
DB 1441 GCATTGTCAAAGCAGCGTGCATGCTGAAAAATCATATGAGGGCTTGTGGATTGGCC 1500
OY 1554 CCAGAGAAATACATGATTTCTCGAGTACAGATGA 1586
DB 1501 CCAGAGAAATACATGATTTCTCGAGTACAGATGA 1533

RESULT 11
ADS82005
ID ADS82005 standard; cDNA; 1533 BP.
AC ADS82005;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
XX
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
XX
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
XX
PA (SEBA/) SEBASTIAN S A.
XX
PI Hitz WD, Sebastian SA;
XX
XX
XX MPI: 2004-639957/62.
XX
XX P-PSDB; ADS82006.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
```

PT phylic acid and inorganic phosphate content of soybean seeds.
XX
XX
PS Claim 2; SEQ ID NO 15; 34pp; English.
XX

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phylic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate), and making or producing a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phylic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

XX
XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

XX
XX Query Match 83.3%; Score 1465.8; DB 13; Length 1533;

XX
XX Best Local Similarity 97.3%; Pred. No. 0;

XX
XX Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
OY 54 ATGTTTCATCGAGATTTTAAAGTTGAGTGTCTTATGTGAGTACACCGAGCTGAGAT 113
DB 1 ATGTTTCATCGAGATTTTAAAGTTGAGTGTCTTATGTGAGTACACCGAGCTGAGAT 60
OY 114 CAGTCCGCTGTAACCTAGCAACCAACCGAATCTGTTTCAAGAGACGAGATGGACCTAT 173
DB 61 CAGTCCGCTGTAACCTAGCAACCAACCGAATCTGTTTCAAGAGACGAGATGGACCTAT 120
OY 174 CAGTGGATTTGCAAAACCCAAATCTGTCAATACGAATTTAAACCAACATCCATGTTCT 233
DB 121 CAGTGGATTTGCAAAACCCAAATCTGTCAATACGAATTTAAACCAACATCCATGTTCA 180
OY 234 AAATTAGGGGTAATGCTTGTGGGTGGGTGAAAACAACGCTCAACCTCAACCGGTGT 293
DB 181 AAATTAGGGGTAATGCTTGTGGGTGGGTGAAAACAACGCTCAACCTCAACCGGTGT 240
OY 294 GTTATTGCTTAACGAGAGGGCATTTTCATGAGGCTAACAAAGACAGATTTCAACAGCAAT 353
DB 241 GTTATTGCTTAACGAGAGGGCATTTTCATGAGGCTAACAAAGACAGATTTCAACAGCAAT 300
OY 354 TACTTTGGCTCCCTCAACCCAAACCTCAGTATCCGATGTGGGTCTTCCAGGAGAGGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAAACCTCAGTATTCGAGTTGATCTCTCCAGGAGAGGAA 360
OY 414 ATCTATGCCCATTTCAAGAGCTGCTTCCATGTTTAACTCTGACGACATTTGTGTTGGG 473
DB 361 ATCTATGCCCATTTCAAGAGCTGCTTCCATGTTTAACTCTGACGACATTTGTGTTGGG 420
OY 474 GGATGGATATAGCAACATGAACCTGCTGATGTCATGCGACAGGGCAAAAGGTTTGAC 533
DB 421 GGATGGATATAGCAACATGAACCTGCTGATGTCATGCGACAGGGCAAAAGGTTTGAC 480
```



```

QY 534 ATGATTTGAGAGCAGTTGAGGCTTACATGAAATCCATGCTCACTCCCGGAATC 593
DB 481 ATGATTTGAGAGCAGTTGAGGCTTACATGAAATCCATGCTCACTCCCGGAATC 540
QY 594 TATGACCCGGATTTCATTTGCTGCAACCAAGAGAGGCTCCAAACAATCATCAAGGCG 653
DB 541 TAGGACCCGGATTTCATTTGCTGCAACCAAGAGAGGCTCCAAACAATCATCAAGGCG 600
QY 654 ACAAGCAAGAGAGGTTTCAACAATCATCAAGAGAGGCTCCAAACAATCATCAAGGCG 713
DB 601 ACAAGCAAGAGAGGTTTCAACAATCATCAAGAGAGGCTCCAAACAATCATCAAGGCG 660
QY 714 AAGTGAACAAGGTTGTTGCTGCTGCAACAAGAGAGGTTTCAACAATCATCAAGGCG 773
DB 661 AAGTGAACAAGGTTGTTGCTGCTGCAACAAGAGAGGTTTCAACAATCATCAAGGCG 720
QY 774 GTGGGCTTAAATGACACATGAGAAATCTCTGCTGCTGCAACAAGAGAGGTTTCAAG 833
DB 721 GTGGGCTTAAATGACACATGAGAAATCTCTGCTGCTGCAACAAGAGAGGTTTCAAG 780
QY 834 ATTTCTCTTCCACCTTTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 893
DB 781 ATTTCTCTTCCACCTTTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 894 GGAAGCCTCAGAACCTTTTGAACAAGGCTGATTTGATTTGCTGCTGCTGCTGCTGCT 953
DB 841 GGAAGCCTCAGAACCTTTTGAACAAGGCTGATTTGATTTGCTGCTGCTGCTGCTGCT 900
QY 954 TTGATTTGAGAGATGATTTCAAGAGTGTGCAAGCAACAATGAAATCTGTTGTTGAT 1013
DB 901 TTGATTTGAGAGATGATTTCAAGAGTGTGCAAGCAACAATGAAATCTGTTGTTGAT 960
QY 1014 TTCTTTGAGGAGGCTGTGATCAAGCAACAATGATTTGATTTGATTTGATTTGAT 1073
DB 961 TTCTTTGAGGAGGCTGTGATCAAGCAACAATGATTTGATTTGATTTGATTTGAT 1020
QY 1074 AATGATGATGATGATTTTGGGCTCCAAACAATTTCCGTTCCAAAGAAATCTCCAAAG 1133
DB 1021 AATGATGATGATGATTTTGGGCTCCAAACAATTTCCGTTCCAAAGAAATCTCCAAAG 1080
QY 1134 AACGTTTGTAGATGATTTGATCAAGCAACAATGCTGCTGATGAGCTGCAATCCA 1193
DB 1081 AACGTTTGTAGATGATTTGATCAAGCAACAATGCTGCTGATGAGCTGCAATCCA 1140
QY 1194 GACATGTTGTTGTTTAAATGATGCTTAAAGAGGAGCAAGCAAGAGGCTGAT 1253
DB 1141 GACATGTTGTTGTTTAAATGATGCTTAAAGAGGAGCAAGCAAGAGGCTGAT 1200
QY 1254 GAGTACCTTCAAGATTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1313
DB 1201 GAGTACCTTCAAGATTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
QY 1314 GAGGATTTCCCTTAACTGCTCTTAAATGATGCTTAAAGAGGAGGAGGAGGAGGAGG 1373
DB 1261 GAGGATTTCCCTTAACTGCTCTTAAATGATGCTTAAAGAGGAGGAGGAGGAGGAGG 1320
QY 1374 ACTAGATGAGTTTAAAGCTGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1433
DB 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
QY 1434 ACCATCTGAGCTCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1493
DB 1381 ACCATCTGAGCTCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
QY 1494 GATTTGTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1553
DB 1441 GATTTGTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
QY 1554 CCAGAGATTAACATGATTTCTCGATCAAGTGA 1586
DB 1501 CCAGAGATTAACATGATTTCTCGATCAAGTGA 1533

```

```

RESULT 12
ADSS82003
ID ADSS82003 standard; cDNA, 1533 BP.
XX
AC ADSS82003;
XX
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; line 29018BP03.
XX
OS
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
XX
PR 07-APR-1998; 98WO-US006822.
XX
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX
PI Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
XX
XX P-PSDB; ADS82004.
XX
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phylic acid and inorganic phosphate content of soybean seeds.
XX
XX
PS
XX
XX Example 8; SEQ ID NO 13; 34pp; English.
XX
XX
CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phylic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homologous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phylic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

```


PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX HITZ WD, Sebastian SA, Grace DJ, Streit LG;
XX MPI; 2004-533135/51.
XX P-PSDB; AD04501.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX PT useful for producing plants with decreased raffinose, stachyose, and
XX PT phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Claim 10; SEQ ID NO 11; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to a suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents cDNA encoding a mutant soybean
XX myo-inositol 1-phosphate synthase polypeptide of the invention.
SQ
Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other:
Query Match 83.2%; Score 1464.2; DB 12; Length 1533;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 54 ATGTTATCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCGAGACTGAGATT 113
DB 1 ATGTTATCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCGAGACTGAGATT 60
QY 114 CAGTCCGTGTACAACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGSCACTAT 173
DB 61 CAGTCCGTGTACAACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGSCACTAT 120
QY 174 CAGTGGATTGTCAAAACCAATTTGCAATACGAATTTTAAACCAACATCCATGTTCT 233
DB 121 CAGTGGATTGTCAAAACCAATTTGCAATACGAATTTTAAACCAACATCCATGTTCT 180
QY 234 AAATTAGGGGTATATGCTTGGGGTGGGAGAAACAGGCTCAACCTCACCGGTGAT 293
DB 181 AAATTAGGGGTATATGCTTGGGGTGGGAGAAACAGGCTCAACCTCACCGGTGAT 240
QY 294 GTTATTTGCTAACGAGAGGCAATTTGATGGGCTACAAAGAACAGAAATTCACAGCAAT 353
DB 241 GTTATTTGCTAACGAGAGGCAATTTGATGGGCTACAAAGAACAGAAATTCACAGCAAT 300
QY 354 TACTTTGGCTCCCTCAACCCAGCTCAGCTATCCGAGTTGGGTCTTTCAGAGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAGCTCAGCTATTCGAGTTGATCTTTCAGAGAGAGAA 360
QY 414 ATCTATGCCCAATTAAAGGCTGCTTCAATGTTTAAACCTGACCACTTTGTTGGG 473
DB 361 ATCTATGCCCAATTAAAGGCTGCTTCAATGTTTAAACCTGACCACTTTGTTGGG 420
QY 474 GGATGGGATATGCAACATGAACCTGGCTGATGCCATGGCGCAAGAGGTGTTGAC 533
DB 421 GGATGGGATATGCAACATGAACCTGGCTGATGCCATGGCGCAAGAGGTGTTGAC 480

QY 534 ATGATTTTGCAGAGAGGTTGAGGCTTACATGGAATTCATGCTTCCACTCCCGGAATC 593
DB 481 ATGATTTTGCAGAGAGGTTGAGGCTTACATGGAATTCATGCTTCCACTCCCGGAATC 540
QY 594 TATGACCCGGAATTTTCAATGCTGCAACCAAGAGAGCGTGCACCAACAGCTCATCAAGGC 653
DB 541 TATGACCCGGAATTTTCAATGCTGCAACCAAGAGAGCGTGCACCAACAGCTCATCAAGGC 600
QY 654 ACAAGACAGAGAGGTTTCAATCATCAAAAGACATCAAGCGTTTAAAGAACCAACC 713
DB 601 ACAAGACAGAGAGGTTTCAATCATCAAAAGACATCAAGCGTTTAAAGAACCAACC 660
QY 714 AAAGTGGAAGGTTGTTGATCTGTGACCTGCCAACACAGAGGATCAATTTGGTT 773
DB 661 AAAGTGGAAGGTTGTTGATCTGTGACCTGCCAACACAGAGGATCAATTTGGTT 720
QY 774 GTGGGCTTAAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGCTGAG 833
DB 721 GTAGGCTTAAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGCTGAG 780
QY 834 ATTTCTCTTCCACTTGTATGCAATGCTGTGTGTATGAAATGTTCTTTCAATTAAT 893
DB 781 ATTTCTCTTCCACTTGTATGCAATGCTGTGTGTATGAAATGTTCTTTCAATTAAT 840
QY 894 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATGCCAGAACACT 953
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATGCCAGAACACT 900
QY 954 TTGATTTGAGGAGATGACTTCAAGAGTGTCTGACACCAAAATGAATCTGTGTTGGTAT 1013
DB 901 TTGATTTGAGGAGATGACTTCAAGAGTGTCTGACACCAAAATGAATCTGTGTTGGTAT 960
QY 1014 TTCTTTGGGGGCTGTGTATCAAGCCCAATCATATAGTCAATCAACCATCTGGGAAAC 1073
DB 961 TTCTTTGGGGGCTGTGTATCAAGCCCAATCATATAGTCAATCAACCATCTGGGAAAC 1020
QY 1074 AATGATGATATGAATCTTTGGCTCCACAACTTTCCGTTCCAGAAATCTTCCAGAAC 1133
DB 1021 AATGATGATATGAATCTTTGGCTCCACAACTTTCCGTTCCAGAAATCTTCCAGAAC 1080
QY 1134 AAGGTTGTGATGATATGCTTCAACAGCAATGCAATCTCTATAGAGCTGTGTAATCCA 1193
DB 1081 AAGGTTGTGATGATATGCTTCAACAGCAATGCAATCTCTATAGAGCTGTGTAATCCA 1140
QY 1194 GACATGTTGTTGTTATTAATGATGCTTCAAGGAGACAGCAAGAGCCATGAT 1253
DB 1141 GACATGTTGTTGTTATTAATGATGCTTCAAGGAGATGCAAGAGAGCCATGAT 1200
QY 1254 GAGTACACTTCAAGATATTCATGGGTGAAGAGACCAATGTTTTCACACATGC 1313
DB 1201 GAGTACACTTCAAGATATTCATGGGTGAAGAGACCAATGTTTTCACACATGC 1260
QY 1314 GAGGATTCCTCTGAGCTCTCTATTAATCTTGAATGCTGTCTTCTGAGCTCAGC 1373
DB 1261 GAGGATTCCTCTTGAAGCTCTCTATTAATCTTGAATGCTGTCTTCTGAGCTCAGC 1320
QY 1374 ACTAGATTCGAGTTTAAAGCTGAAATGAGGAAATTCACATCTTCCACCAAGTTGCT 1433
DB 1321 ACTAGATTCGAGTTTAAAGCTGAAATGAGGAAATTCACATCTTCCACCAAGTTGCT 1380
QY 1434 ACCATCTCAGCTTACCTCAACAGGCTCTCTGTTTCCACGGGATCACCAAGTGTGAT 1493
DB 1381 ACCATCTCAGCTTACCTCAACAGGCTCTCTGTTTCCACGGGATCACCAAGTGTGAT 1440
QY 1494 GCAATTTGCAAGCAGCGTGCATATGCTGGAACCAATTAAGAGGCTTGTGGAATGGCC 1553
DB 1441 GCAATTTGCAAGCAGCGTGCATATGCTGGAACCAATTAAGAGGCTTGTGGAATGGCC 1500
QY 1554 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1586
DB 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 14
ADS82001
ID ADS82001 standard; cDNA; 1533 BP.
XX ADS82001;
AC
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
OS Glycine max; line 29010CP01.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
FT replace(260,G)
FT mutation /tag= b
FT
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI: 2004-639957/62.
XX P-PsDB; ADS82002.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 11; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful

CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a mutant myo-
CC inositol 1-phosphate synthase.
XX

SQ Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 83.2%; Score 1464.2; DB 13; Length 1533;

Best Local Similarity 97.2%; Pred. No. 0;

Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 54 ATGTTTCATCCGAAATTTTAAAGTGTGAGTGTCTTAATGGAATGACCGAGATCGAGTT 113
DB 1 ATGTTTCATCCGAAATTTTAAAGTGTGAGTGTCTTAATGGAATGACCGAGATCGAGTT 60
QY 114 CAGTCCGTTGTCACAACTACGAAACCCGAACTGTTTCAGAGAAACAGAAATGACCTAT 173
DB 61 CAGTCCGTTGTCACAACTACGAAACCCGAACTGTTTCAGAGAAACAGAAATGACCTAT 120
QY 174 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAACATCTGTTCT 233
DB 121 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAACATCTGTTCT 180
QY 234 AAATTAGGGGTAAATGCTTGTGGGTGGGTGGGAACAAACGGCTCAACCTTCAACGGTGT 293
DB 181 AAATTAGGGGTAAATGCTTGTGGGTGGGTGGGAACAAACGGCTTCAACCTTCAACGGTGT 240
QY 294 GTTATTTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGACCAAGATTCAACACCAAT 353
DB 241 GTTATTTGCTAACCGAGAGGACATTTTCATGGGCTACAAAGACCAAGATTCAACACCAAT 300
QY 354 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 360
QY 414 ATCTATGCCCCCATTTAAAGAGCTGTTTCCAAATGTTAACCTTGACGACATTTGTTGGG 473
DB 361 ATCTATGCCCCCATTTAAAGAGCTGTTTCCAAATGTTAACCTTGACGACATTTGTTGGG 420
QY 474 GGATGGATATTCAGCAATGAACTGAGCTATGCTGACGACGAGGCAAAAGTGTGTTGAC 533
DB 421 GGATGGATATTCAGCAATGAACTGAGCTATGCTGACGACGAGGCAAAAGTGTGTTGAC 480
QY 534 ATCGATTTGCAAGAACAGTTGAGGCTTTACATGAAATCATGATCTTCCATCTCCCGGAATC 593
DB 481 ATCGATTTGCAAGAACAGTTGAGGCTTTACATGAAATCATGATCTTCCATCTCCCGGAATC 540
QY 594 TATGACCCCGGATTTTATTTGCTGCTCCCAACCAAGAGAGCTGCCAACAAGTATCAAGGGC 653
DB 541 TACGACCCCGGATTTTATTTGCTGCTCCCAACCAAGAGAGCTGCCAACAAGTATCAAGGGC 600
QY 654 ACAAGCAAGAGCAAGTTCAACAAATCATCAAAAGACATCAAGGCGTTTAAAGAACCCACC 713
DB 601 ACAAGCAAGAGCAAGTTCAACAAATCATCAAAAGACATCAAGGCGTTTAAAGAACCCACC 660
QY 714 AAAGTGCACAAAGTGTGTTGACTGTGACCTGTCACACAGAGAGGTACAGTAATTTGGTT 773
DB 661 AAAGTGCACAAAGTGTGTTGACTGTGACCTGTCACACAGAGAGGTATAGCAATTTGGTT 720
QY 774 GTGGGCTTTAATGACACATGAGAAATCTTTGGTCTGCTGTGTGACAGAAATGAGAGCTGAG 833
DB 721 GTAGGCTTTAATGACACATGAGAAATCTTTGGTCTGCTGTGTGACAGAAATGAGAGCTGAG 780
QY 834 ATTTCCTCCTCCACTTGTATGAGCCATGCTGTGTGTTAGAAATGTTCTTTCAATAT 893
DB 781 ATTTCCTCCTCCACTTGTATGAGCCATGCTGTGTGTTAGAAATGTTCTTTCAATAT 840
QY 894 GGAAGCCCTCAGAACACTTTGTTGACAGAGGCTGATTTGATCTTCCATCGCAGAGAACCT 953
DB 841 GGAAGCCCTCAGAACACTTTGTTGACAGAGGCTGATTTGATCTTCCATCGCAGAGAACCT 900
QY 954 TTGATTTGGTGAGATGACTTCAAGAGTGTGACACCAAAATGAATCTGTGTGGTTGAT 1013

DB 901 TTGATGTGTGAGATGACTTCAAGAGTGTGACAGACCAAAATGAAATCTGNTGTTGAT 960
QY 1014 TTCTTGTGGGGCTGGTATCAAGCCAACTATATGTAGTACATTAACCATCTGGGAAAC 1073
XX |||||
CC 961 TTTCTGTGGGGCTGGTATCAAGCCAACTATATGTAGTACATTAACCATCTGGGAAAC 1020
DB |||||
QY 1074 AATGATGTATGAAATCTTCCGGCTCCAAACAACTTCCGTTCCAAAGAAATCTCCAGAGC 1133
DB 1021 AATGATGTATGAAATCTTCCGGCTCCAAACAACTTCCGTTCCAAAGAAATCTCCAGAGC 1080
QY 1134 AACGTTGTATGATATGTGTCAACAGCAATGCCATCTCTATAGAGCTGTGAAATCCA 1193
DB 1081 AACGTTGTATGATATGTGTCAACAGCAATGCCATCTCTATAGAGCTGTGAAATCCA 1140
QY 1194 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGGAGACAGCAAGAGCCATGAT 1253
DB 1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGGAGATAGCAAGAGCCATGAT 1200
QY 1254 GAGTACACTTCAAGATATTCATGGGTGAAAGAGACCAATGTTTGGACAAACATATGC 1313
DB 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAGACCAATGTTTGGACAAACATATGC 1260
QY 1314 GAGATTCCTCTTGTAGCTCTCCTATTATCTTGAATGATCTTCTGTGAGCTCAGC 1373
DB 1261 GAGATTCCTCTTGTAGCTCTCCTATTATCTTGAATGATCTTCTGTGAGCTCAGC 1320
QY 1374 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTCATCTCATTCACCCAGTTGCT 1433
DB 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTCATCTCATTCACCCAGTTGCT 1380
QY 1434 ACCATCTCAGCTACTCAACCAAGGCTCTCTGTTCCACCGGTACACAGGTGAT 1493
DB 1381 ACCATCTCAGCTACTCAACCAAGGCTCTCTGTTCCACCGGTACACAGGTGAT 1440
QY 1494 GCATTGTCAAGCAGCGTCAATGCTGAAACCATATGAGGCTGTGTTGATGGCC 1553
DB 1441 GCATTGTCAAGCAGCGTCAATGCTGAAACCATATGAGGCTGTGTTGATGGCC 1500
QY 1554 CCAGAGATTAACATGATTTCTCGATACAGTGA 1586
DB 1501 CCAGAGATTAACATGATTTCTCGATACAGTGA 1533

RESULT 15
AAK90402
ID AAK90402 standard; cDNA to mRNA; 1950 BP.
AC AAK90402;
XX
XX 24-SEP-1999 (first entry)
DB Nicotiana paniculata INPS encoding cDNA.
XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
KW water stress; resistance; de.
XX Nicotiana paniculata.
OS
XX
XX JPI1187879-A.
PN
XX
XX 13-JUL-1999.
PD
XX
XX 26-DEC-1997; 97JP-00359773.
PF
XX
XX 26-DEC-1997; 97JP-00359773.
PR
XX
XX (NISB) JAPAN TOBACCO INC.
PA
XX
XX WPI; 1999-451546/38.
DR P-PSDB; AAY24477.
XX
XX New INPS gene derived from Nicotiana genus plant - useful for conferring
PT resistance to water stress to plants.

XX
PS Claim 2; Page 6-8; bpp; Japanese.
XX
CC The present sequence encodes Nicotiana paniculata inositol monophosphate
CC synthase (INPS), designated NpINPS1. INPS can be used to confer water
CC stress resistance to a plant
XX
SQ Sequence 1950 BP; 570 A; 377 C; 425 G; 578 T; 0 U; 0 Other;
Query Match 63.7%; Score 1120.4; DB 2; Length 1950;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 261; Indels 0; Gaps 0;
QY 51 ATATATGTTATGAGATTTTAAAGTTGAGTGTCTTAATGTGAATGACCCGAGCTGAG 110
DB 89 AAAATGTTATTTAGAAATTTTAAAGTTGAGTGTGAGCCCAAGTTAAGTACCAAAAGTGA 148
QY 111 ATTCAGTCCGTGTACAACTACGAAACCAAGAACTGTTACAGAGAACAGATGACCC 170
DB 149 ATTCAGTCTGTCTATGATTTATCAAAACCTGAGTTAGTTATGATGAGAAATGGAGCA 208
QY 171 TATCAGTGTATGTCAAAACCAATCTGTCAAAATGAAATTTAAACCAACATCCATGTT 230
DB 209 TATCAATGAGCCGTCAAGCTTAAGACTGTCAAAATGAGTTCAAGCTGATGTTATGTT 268
QY 231 CCTAAATTTAGGGTAAATGCTTGTGGTTGGGTGAAACACAGGCTCAACCTCACCCT 290
DB 269 CCCAAATTTAGGGGTAAATGCTTGTGGTTGGGTGAAACAAATGTTCAACCTTGACCGT 328
QY 291 GGTGTTATGCTTACCGAGAGGGCAATTCATGGGCTACAAAGAGCAAGATTCACAGGCC 350
DB 329 GGTGTTATGCTTACCGAGAGGGCAATTCATGGGCTACCGAGAGCAAGATTCACAGGCC 388
QY 351 AATTAATTTGAGCTCTCCCTCAACCAAGCTTCACTATCCAGTTGGGTCTTTCAGAGGAG 410
DB 389 AATTAATTTGAGCTCTCTTCACTCAAGGCTTCACTATCCAGTTGGGTCTTTCAGAGGAG 448
QY 411 GAAATCTATGCCCCATTCAAGAGCTGCTTCCAAATGTTTAAACCTGACGACATGTGTT 470
DB 449 GAGATCTATGCTCAATTTAAAGCTCTCTTCCAAATGTTTAAACCTGACGACATGTGTT 508
QY 471 GGGGAGTGGGATATGAGCAACATGAACCTGCTGATGTCATGAGCCAGAGGCAAGGCTTT 530
DB 509 GAGAGTGGGATATGAGCAACATGAACCTGCTGATGTCATGAGCCAGAGGCAAGGCTTT 568
QY 531 GACATGATTTGACAGAGCAGTTGAGGCTTACATGATCCATGCTTCCATCCCGGAG 590
DB 569 GATATTGATCTTCAAAAGAGTTGAGGCTTCAATGATCCATGCTTCCATCCCGGAG 628
QY 591 ATCTATGACCCGAGATTTCAATGCTGCTGCCAACAAAGAGGCTGCCAACACGTCATCAG 650
DB 629 ATCTATGACCCGAGATTTCAATGCTGCTGCCAACAAAGGCTGCCAACACGTCATCAG 688
QY 651 GGCACAAAGCAAGACCAATGCTCAACATCAACAAAGCATCAAGGCTTTAAGAGAGCC 710
DB 689 GGAACCAAGAAAGCAAAATGATCAAAATCAATTAAGATTAAGAGAGTTAAGAGAGAG 748
QY 711 ACCAAAGTGAACAAGGTGTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 770
DB 749 AACCAAGTGAACAAGGTGTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 808
QY 771 GTTGTGGGCTTTAATGACCAATGAGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 830
DB 809 GTTGTGGGCTTTAATGACCAATGAGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 868
QY 831 GAGATTTCTCTTCAACCTGTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 890
DB 869 GAAATATCTCTTCAACCTGTGATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
QY 891 AATGAGCCCTCAGAACATTTTGTACAGGCTGATGATCTTGTGACGAGAGC 950
DB 929 AATGAGCCCTCAGAACATTTTGTACAGGCTGATGATCTTGTGACGAGAGC 988

```
OY 951 ACTTGATTGGTGGAGATGACTTCAAGATGGTGCAGACCACCAAAATGAAATCTGTGTGGTT 1010
    |||||
Db 989 ACATTGATTGGTGGTGGATGACTTTAAAGTGGTCAACCAAAATGAGTCAGTGGT 1048
    |||||
OY 1011 GATTTCTTTGGGGGCTGGTATCAAGCCACATCTATAGTCAGTTACACCATCTGGGA 1070
    |||||
Db 1049 GATTTCTTTGGAGCTGGTATTAAGCCACATCATATGTGAGCTACACCATTTGGGT 1108
    |||||
OY 1071 AACATGATGTGATTAATCTTCCGCTCCACAACTTCCGTTCCAGAGAAATCTCCAG 1130
    |||||
Db 1109 AACAAATGATGATTAATCTTCCGCTCCACAACTTCCGTTCCAGAGAAATCTCCAG 1168
    |||||
OY 1131 AGCAAGTGTGATGATGATGATGATCAACAGCAATGCCATCTCTATAGGCTGTGAACAT 1190
    |||||
Db 1169 AGTAATGTTGTGATGATGATGATGATCAACAGCAATGCCATCTCTATAGGCTGTGAACAT 1228
    |||||
OY 1191 CCAGACCATGTTGTTGTTATTAAGTATGTCCTTAAGTATGGGACACAGAGAGCATG 1250
    |||||
Db 1229 CCGACCATGTTGTTGTTATTAAGTATGTCCTTAAGTATGGGACACAGAGAGCATG 1288
    |||||
OY 1251 GATGAGTACACTTCAGAGATATTCATGGGTGGAAGAGACCATGTTTGCACAAACAC 1310
    |||||
Db 1289 GATGAGTACACTTCAGAGATATTCATGGGTTGGAAGAGACCATGTTTGCACAAACAT 1348
    |||||
OY 1311 TGCAGAGATTCCTCTTAAGCTGCTCTATATCTTGACTGTGCTTCTTGTGAGCTC 1370
    |||||
Db 1349 TGTGAGATTCCTCTTAAAGCTGCTCTCAATATATGATTTGTGCTTCTTGTGAGCTC 1408
    |||||
OY 1371 AGCACTAGATGAGTTTAAAGCTGAAATGAGGAAATTCCTCATTCATCCACCGATT 1430
    |||||
Db 1409 AGTACCCGATTCAGCTCAAGCTGAAAGAGAGGATGATTCCTCATTCACCCGATT 1468
    |||||
OY 1431 GGTACCATCTCAAGCTCACTCAACAGGCTCTCTGTGTCCACCGGTTACACAGTGTG 1490
    |||||
Db 1469 GGTACTATCTCAAGCTCACTCAACAGGCTCTCTGTGTACACAGGTTACACAGTGTG 1528
    |||||
OY 1491 AATGATTTGCAAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATG 1550
    |||||
Db 1529 AATGATTTGCAAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATG 1588
    |||||
OY 1551 GCCCGAGAGATTAACATGATTTCTCGAGTACAAATGAAG 1588
    |||||
Db 1589 GCACCGAGAGATTAACATGATTTCTCGAGTACAAATGAAG 1626
    |||||
```

Search completed: June 7, 2005, 22:54:21
Job time : 981.454 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 20:51:42 ; Search time 7677.08 Seconds
(without alignment)
1108.537 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctctatctatctcttctgt.....aatgcagataattctgt 1760

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapect 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenBank1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1760	100.0	1782	6	BD075266 Soybean p
2	1750.4	99.5	1781	8	AF293970 Glycine m
3	1717	97.6	1739	8	AY038802 Glycine m
4	1531.4	87.0	1533	6	BD075269 Soybean p
5	1120.4	63.7	1931	8	AB032073 Nicotiana
6	1120.4	63.7	1950	6	E27176 Novel INPS
7	1115.6	63.4	1959	8	AB009881 Nicotiana
8	1092.2	62.1	1954	8	BT013759 Lycopersi
9	1087	61.8	1959	8	BT013505 Lycopersi
10	1073.4	61.0	1978	8	CPINOIG Z32632 C.pardisi
11	1064.2	60.5	1845	8	AF284065 Sesamum 1
12	1037.8	59.0	2053	8	MCU32511 Mesembryant
13	1037.4	58.9	1986	8	AY028259 Aricennia
14	1010.8	57.4	1863	8	AF433879 Suaeda ma
15	1008.4	57.2	1863	8	AY065415 Arabidops
16	1006.8	57.2	1863	8	AY096554 Arabidops
17	1006.8	57.2	1864	8	AY054202 Arabidops
18	1006.8	57.2	1864	8	AY053415 Arabidops
19	1002	56.9	1890	8	ATU30250 Arabidops

20	1001.4	56.9	1781	6	AX054630 Sequence
21	1001.4	56.9	1781	8	BN066307 U66307 Brsescia na
22	999.6	56.8	1538	8	AF120146 AF120146 Triticum
23	999.6	56.8	1538	8	AF120147 AF120147 Triticum
24	999.6	56.8	1538	8	AF120148 AF120148 Triticum
25	999.2	56.8	1914	8	AK103501 Oryza sat
26	999.2	56.8	1915	8	AK058750 Oryza sat
27	997.6	56.7	1868	8	AB012107 Oryza sat
28	996.2	56.6	1870	8	AF372954 Arabidops
29	995.4	56.6	1533	6	CQ080508 Sequence
30	995.4	56.6	1533	6	CQ080508 Sequence
31	995.4	56.6	1533	8	AY143904 Arabidops
32	995.4	56.6	1533	8	AY143904 Arabidops
33	994.6	56.5	1837	8	AY085931 Arabidops
34	994.2	56.5	1859	6	AX151207 Sequence
35	990.8	56.3	1567	8	BT001931 Arabidops
36	990.8	56.3	1665	6	BD073470 Regulated
37	990.8	56.3	1665	6	AF056326 Zea mays
38	990.4	56.3	1921	8	ATU04876 Arabidops
39	988.8	56.2	1871	8	PVU38920 Phaseolus v
40	987.4	56.1	1535	6	BD073472 Regulated
41	985.6	56.0	2152	8	AF056325 Hordeum v
42	984.6	55.9	1936	8	AB059557 Avena sat
43	984	55.9	1931	6	AR137808 Sequence
44	984	55.9	1931	6	AR170167 Sequence
45	982.2	55.8	1930	8	AY035045 Arabidops

ALIGNMENTS

RESULT 1	BD075266	1782 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD075266				
DEFINITION	Soybean plant producing seeds with reduced levels of raffinose				
ACCESSION	BD075266.1	GI:22620869			
VERSION	JP 2001519665-A/1.				
KEYWORDS	unclassified				
SOURCE	unclassified				
ORGANISM	unclassified				
REFERENCE	1 (bases 1 to 1782)				
AUTHORS	Hitz,W.D. and Sebastian,S.A.				
TITLE	Soybean plant producing seeds with reduced levels of raffinose				
JOURNAL	patent: JP 2001519665-A 1 23-OCT-2001;				
COMMENT	SI DU PONT DE NEMOURS AND CO				
	OS Soybean line LR13				
	PN JP 2001519665-A/1				
	PD 23-OCT-2001				
	PF 07-APR-1998 JP 1998543012				
	PR 08-APR-1997 US 08/835751				
	PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN				
	PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC				
	Strandedness: Double;				
	CC Topology: Linear;				
	CC Soybean plant producing seeds with reduced				
	levels of raffinose				
	CC and phytic acid				
	FM Key	Location/Qualifiers			
	FT CDS	54..1586.			
FEATURES	source	Location/Qualifiers			
		1..1782			
		/organism="unclassified"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:32644"			
ORIGIN	Query Match	100.0%; Score 1760; DB 6; Length 1782;			
	Best Local Similarity	100.0%; Pred. No. 0;			
	Matches 1760; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			

1 CTTCTTTATTCCTTTTGTATTCATTCATTCCTTAATCTTGTGAAAAATATGTTCA 60
 1 CTTCTTTATTCCTTTTGTATTCATTCATTCCTTAATCTTGTGAAAAATATGTTCA 60
 61 TCGAATAATTTAAGGTGAGTGTCTTAATGTAGATACCGAGACTGAGATTCAGTCCG 120
 61 TCGAATAATTTAAGGTGAGTGTCTTAATGTAGATACCGAGACTGAGATTCAGTCCG 120
 121 TGTACAACTACGAAACCAACCGAACTTTGTACAGAGAACGGAATGGCACTTCAGTGA 180
 121 TGTACAACTACGAAACCAACCGAACTTTGTACAGAGAACGGAATGGCACTTCAGTGA 180
 181 TGTGCAAAACCAATCTGTCAATACGAATTTAAAAACCAACATTCATGTTCTTAATTAG 240
 181 TGTGCAAAACCAATCTGTCAATACGAATTTAAAAACCAACATTCATGTTCTTAATTAG 240
 241 GGGTAATGCTTGTGGGTTGGGGTGGAAACAAACGGCTCAACCTCAACCGGTGTGTTATTG 300
 241 GGGTAATGCTTGTGGGTTGGGGTGGAAACAAACGGCTCAACCTCAACCGGTGTGTTATTG 300
 301 CTAAACCGAAGGGGCAATTTATGGGCTACAAAGAACAAATTCAAAGCCATTTACTTTG 360
 301 CTAAACCGAAGGGGCAATTTATGGGCTACAAAGAACAAATTCAAAGCCATTTACTTTG 360
 361 GCTCCCTCAACCAAGGCTCAGCTATCCGAGTTGGGCTCTCCAGGGAGAGAAATCTATG 420
 361 GCTCCCTCAACCAAGGCTCAGCTATCCGAGTTGGGCTCTCCAGGGAGAGAAATCTATG 420
 421 CCCCATTCAGAGCCCTGCTTCCAAATGTAAACCTGTACGACATTTGTTGGGGAGATGG 480
 421 CCCCATTCAGAGCCCTGCTTCCAAATGTAAACCTGTACGACATTTGTTGGGGAGATGG 480
 481 AATATAGCAACATGAAACCTGTGCTGATGCTGATGCTGCAAGGGCAAGGTGTTGACATGAT 540
 481 AATATAGCAACATGAAACCTGTGCTGATGCTGATGCTGCAAGGGCAAGGTGTTGACATGAT 540
 541 TSCAAGAGAGTGTGAGGCTTACATGATGATTCATGCTCCCGGAACTATATGACC 600
 541 TSCAAGAGAGTGTGAGGCTTACATGATGATTCATGCTCCCGGAACTATATGACC 600
 551 TCGAAGAGAGTGTGAGGCTTACATGATGATTCATGCTCCCGGAACTATATGACC 600
 551 TCGAAGAGAGTGTGAGGCTTACATGATGATTCATGCTCCCGGAACTATATGACC 600
 601 CGGATTTCAATGCTGTGCAACCAAGAGAGCGTGCCAAACGTCATCAAGGGGCAACAAAGC 660
 601 CGGATTTCAATGCTGTGCAACCAAGAGAGCGTGCCAAACGTCATCAAGGGGCAACAAAGC 660
 661 AAGAGCAATTCGAACAAATCATCAAGAGATCAAGGCGTTTAAAGGAAGCACCAAGTGG 720
 661 AAGAGCAATTCGAACAAATCATCAAGAGATCAAGGCGTTTAAAGGAAGCACCAAGTGG 720
 721 ACAAGGTGTTGATCTGTGATGCTGCAACAGAGAGTATACGTAATTTGTTGTGGGCC 780
 721 ACAAGGTGTTGATCTGTGATGCTGCAACAGAGAGTATACGTAATTTGTTGTGGGCC 780
 781 TTAATGACACCAATGAGAAATCTTGTGGCTGTGTGACGAATAAGAGCTGAGATTTCTC 840
 781 TTAATGACACCAATGAGAAATCTTGTGGCTGTGTGACGAATAAGAGCTGAGATTTCTC 840
 841 CTTCCACCTTTATATGATGCTGTGTGTTATGGAATAATTTCTTTCATTAATGAAAGC 900
 841 CTTCCACCTTTATATGATGCTGTGTGTTATGGAATAATTTCTTTCATTAATGAAAGC 900
 901 CTCAGAACACTTTTGTACAGGGCTGATGATCTTGTGCAATGCGAGAGAACATTTGATTTG 960
 901 CTCAGAACACTTTTGTACAGGGCTGATGATCTTGTGCAATGCGAGAGAACATTTGATTTG 960
 961 GTGAGATGATCTTCAAGAGTGTGTGACCAACCAATGAAATCTGTGTTGTTGATTTCTTG 1020
 961 GTGAGATGATCTTCAAGAGTGTGTGACCAACCAATGAAATCTGTGTTGTTGATTTCTTG 1020
 1021 TGGGGGCTGTATCAAGCCAACTATATGATGATTAACAACATCTGCGGAAACCAATGATG 1080
 1021 TGGGGGCTGTATCAAGCCAACTATATGATGATTAACAACATCTGCGGAAACCAATGATG 1080

1081 GTATGAATCTTTCGGCTCCACAAACCTTTCGTTCCAGAGAAATCTCCAGAGCAACGTTG 1140
 1081 GTATGAATCTTTCGGCTCCACAAACCTTTCGTTCCAGAGAAATCTCCAGAGCAACGTTG 1140
 1141 TTGATGATATGATGCAACGCAATGCAATCTCTATGAGCCGTGTGAATATCCAGCCATG 1200
 1141 TTGATGATATGATGCAACGCAATGCAATCTCTATGAGCCGTGTGAATATCCAGCCATG 1200
 1201 TTGTTGTTAATTAAGATGATGCTTACGTAGGGGAGCAGAGAGAGCCATGATAGTACA 1260
 1201 TTGTTGTTAATTAAGATGATGCTTACGTAGGGGAGCAGAGAGAGCCATGATAGTACA 1260
 1261 CTTGAGATATATTCATGAGGTGGAAGAGACCATGTTTGTGACAAACATGCGAGATTT 1320
 1261 CTTGAGATATATTCATGAGGTGGAAGAGACCATGTTTGTGACAAACATGCGAGATTT 1320
 1321 CCTCTTAGCTGCTCTTATATCTTGTGAACTGTGCTCTTGTGAGGCTCAGGCACTAAG 1380
 1321 CCTCTTAGCTGCTCTTATATCTTGTGAACTGTGCTCTTGTGAGGCTCAGGCACTAAG 1380
 1381 TCGAGTTTAAAGCTGAATAATGAGGAAATTCACACTCATTCACCCAGTGTCTACATCC 1440
 1381 TCGAGTTTAAAGCTGAATAATGAGGAAATTCACACTCATTCACCCAGTGTCTACATCC 1440
 1441 TCAAGCTACCTCAACCAAGGCTCCTGTGTTCCACCGGATACACAGATGATGATGATGT 1500
 1441 TCAAGCTACCTCAACCAAGGCTCCTGTGTTCCACCGGATACACAGATGATGATGATGT 1500
 1501 CAAAGCAGCGTGCATGCTGTGAAAAATATGAGGGCTTGTGTTGATTTGGCCCAAGAG 1560
 1501 CAAAGCAGCGTGCATGCTGTGAAAAATATGAGGGCTTGTGTTGATTTGGCCCAAGAG 1560
 1561 ATTAATGATTTCTCGAGTACAAAGTGAACATGAGGACCGAAGAAATATATGTTGGGGTAG 1620
 1561 ATTAATGATTTCTCGAGTACAAAGTGAACATGAGGACCGAAGAAATATATGTTGGGGTAG 1620
 1621 CCTAGCTGAATGTTTATGTTAATATATATGTTGCTTAAATTTTGAAGTGAATGAA 1680
 1621 CCTAGCTGAATGTTTATGTTAATATATATGTTGCTTAAATTTTGAAGTGAATGAA 1680
 1681 TGCATCAGCTTCATTAATGCTTTAGAGGGGGCAATCTGTTTATGGAACATGAAATG 1740
 1681 TGCATCAGCTTCATTAATGCTTTAGAGGGGGCAATCTGTTTATGGAACATGAAATG 1740
 1741 AATGATATATATTTTGTGT 1760
 1741 AATGATATATATTTTGTGT 1760

RESULT 2
 AF293970 1791 bp mRNA linear PLN 03-MAY-2001
 LOCUS
 DEFINITION
 cdb.
 Glycine max myo-inositol-3-phosphate synthase (MIRS) mRNA, complete
 ACCESSION
 AF293970
 AF293970.1 GI:13936690
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 1791)
 HEGEMAN, C.E., GOOD, L.L. and GRABAU, E.A.
 Expression of D-myo-inositol-3-phosphate synthase in soybean.
 Implications for phylic acid biosynthesis
 Plant Physiol. 125 (4), 1941-1948 (2001)
 JOURNAL
 MEDLINE
 PUBMED
 1129373
 2 (bases 1 to 1791)
 HEGEMAN, C.E., GOOD, L.L. and GRABAU, E.A.
 Direct Submission
 REFERENCE
 AUTHORS
 TITLE

JOURNAL Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA 24060, USA

FEATURES source Location/Qualifiers

1..1791
/organism="Glycine max"
/mol_type="mRNA"
/cultiivar="Williams 82"
/db_xref="taxon:3847"
/dev_stage="Immature cotyledons"
1..1791
/gene="MIPS"
62..1594
/gene="MIPS"
/note="isomerase"
/codon_start=1
/product="myo-inositol-3-phosphate synthase"
/protein_id="AAK49896.1"
/db_xref="GI:13936691"
/translation="MPIENPKVCEPVKXKTEIOSVNVETTELVHNRNGYQWV
KRSKYREKTNIVHPKGLVGMVGMNGNSTLGTGVIANKRGISWATDOKIQANVF
GSLTQASAIRVGSFQEEIYAPPKSLPWNPDIVFGWDISNMNLADAMARAFD
IDLOQWPMESMLPLPGIYDPDFAANQERANNVIGTQEQVQOIIXDKIAKE
ATKVDKVVLTANTERYSNLVGLNDTMENTLAAVDREABISPTLYAIACMENTV
PRINGSPOFTFVPLGLDIAIARNTLIGDDPKSGOTKMSVLDPLVGAGIKPTISV
VNHLCNNDGMNLSAPOTPRSKISKSNVVDPMVNSAIIYEGEHPDHVVYKYPYV
GDSKRAMDEYTSIFMGKSTIVHAKSTIVLHCTDSLAAFTIIDLVLELSRIEFKANE
GKFHSFHPVATLLSYTKAPLVPFGTGVNALSKOPAMLENTMRACVGLAPENMILE
YK"

ORIGIN

Query Match 99.5%; Score 1750.4; DB 8; Length 1791;
Beet Local Similarity 99.7%; Pred. No. 0;
Matches 1754; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 CTCTCTTATTCCTTTGTAATTTGATTCATTCCTTAATCTTGTGTAATAATATGTTCA 60
9 CTCTCTTATTCCTTTGTAATTTGATTCATTCCTTAATCTTGTGTAATAATATGTTCA 68
61 TCGAGATTTTAAGGTGAGTGTCTTAATGTGAAGTACCGAGACTGAGATTCACTCGG 120
69 TCGAGATTTTAAGGTGAGTGTCTTAATGTGAAGTACCGAGACTGAGATTCACTCGG 128
121 TGTAACTACGAAACCAACCACTGTTCAAGAAACGAAATGGCACTTACATGGA 180
129 TGTAACTACGAAACCAACCACTGTTCAAGAAACGAAATGGCACTTACATGGA 188
181 TGTCAAAACCAAACTGTCAATACGAATTTAAACCAACATCCATGTTCCAAATTG 240
189 TGTCAAAACCAAACTGTCAATACGAATTTAAACCAACATCCATGTTCCAAATTG 248
241 GGGTAATGCTTGTGGGTGGGGTGAACACGGCTCAACCTCAACGGTGGTGTATTG 300
249 GGGTAATGCTTGTGGGTGGGGTGAACACGGCTCAACCTCAACGGTGGTGTATTG 308
301 CTAAACGAGAGGCGATTTATGGGCTACAAAGACAAATTTCAACAAGCCATTTACTTTG 360
309 CTAAACGAGAGGCGATTTATGGGCTACAAAGACAAATTTCAACAAGCCATTTACTTTG 368
361 GGTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGCTCTTCAAGGAGAGGAAATCTTAG 420
369 GGTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGCTCTTCAAGGAGAGGAAATCTTAG 428
421 CCCCATTCAGAGCTGTCTTCAATGTATTACCTGTACGACATTTGTTGGGGAGTGG 480
429 CCCCATTCAGAGCTGTCTTCAATGTATTACCTGTACGACATTTGTTGGGGAGTGG 488
481 AATTACGAACTGAACCTGCTGATGCCATGGCCAGGCGAAAGTGTGTCATGATTT 540
489 AATTACGAACTGAACCTGCTGATGCCATGGCCAGGCGAAAGTGTGTCATGATTT 548
541 TGCAGAACAGTTGAGGCTTACATGAAATCCATGCTTCCCGGGAATCTATGACC 600

549 TCGAGAGCAGTGGAGGCTTACATGGAATCCATGCTTCCACTCCCGGAATCTATGACC 608
601 CGGATTTCTTGTGCGAACCAAGAGAGCTGCGCAACAGTCAATGAGGCGACAAAGC 660
609 CGGATTTCTTGTGCGAACCAAGAGAGCTGCGCAACAGTCAATGAGGCGACAAAGC 668
661 AAGAGCAATTCACAAATCATCAAAAGATCAATGAGGCTTTAAGGAAGCCAAAGTGG 720
669 AAGAGCAATTCACAAATCATCAAAAGATCAATGAGGCTTTAAGGAAGCCAAAGTGG 728
721 ACAAGGTGTTGACTGTGAGCTGCCACACAGAGAGTACAGTAATTTGTTGGGGC 780
729 ACAAGGTGTTGACTGTGAGCTGCCACACAGAGAGTACAGTAATTTGTTGGGGC 788
781 TTAATGACACCATGAGAAATCTTGTGGCTGTGTGACAGAAATGAGCTGAGATTTCTC 840
789 TTAATGACACCATGAGAAATCTTGTGGCTGTGTGACAGAAATGAGCTGAGATTTCTC 848
841 CTTCCACCTGTATGCGCATGCTGTGTGTAAGGAAAAGTTCTCTTCAATTAATGAGGC 900
849 CTTCCACCTGTATGCGCATGCTGTGTGTAAGGAAAAGTTCTCTTCAATTAATGAGGC 908
901 CTCAGAACACTTTGTGACAGAGGCTGATGATCTTGTGCAATGCGAGAGAACATTTGATTG 960
909 CTCAGAACACTTTGTGACAGAGGCTGATGATCTTGTGCAATGCGAGAGAACATTTGATTG 968
961 GTGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGTTGATTTCTTGG 1020
969 GTGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGTTGATTTCTTGG 1028
1021 TGGGGGCTGTATCAAGCCAAATCTATGTCAGTTTACAAACATCTGAGGAAACAAATGAG 1080
1029 TGGGGGCTGTATCAAGCCAAATCTATGTCAGTTTACAAACATCTGAGGAAACAAATGAG 1088
1081 GTATGATCTTTCGGCTCCACAACTTTCGTTCCAAAGAAATCTTCAAGAGCAACGTTG 1140
1089 GTATGATCTTTCGGCTCCACAACTTTCGTTCCAAAGAAATCTTCAAGAGCAACGTTG 1148
1141 TTGATGATATGATCAACAGCAATGCAATCTTATGAGCTGTGTGAACATTCAGACCATG 1200
1149 TTGATGATATGATCAACAGCAATGCAATCTTATGAGCTGTGTGAACATTCAGACCATG 1208
1201 TTGTTGTTTATTAAGTATGTCCTTACGTAGGAGGACAGCAAGAGACCTGATGATGATA 1260
1209 TTGTTGTTTATTAAGTATGTCCTTACGTAGGAGGACAGCAAGAGACCTGATGATGATA 1268
1261 CTTCAAGATATTCATGAGGTGGAAGAGACCATGTTTGTGACAAACATGAGGAGATT 1320
1269 CTTCAAGATATTCATGAGGTGGAAGAGACCATGTTTGTGACAAACATGAGGAGATT 1328
1321 CCCTCTAGCTGCTCTTATTTATCTTGAACCTTGTGCTTGTGCTGAGCTCAGACATGAA 1380
1329 CCCTCTAGCTGCTCTTATTTATCTTGAACCTTGTGCTTGTGCTGAGCTCAGACATGAA 1388
1381 TCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCATCTCATCTCAACCAAGTTGCTACATCC 1440
1389 TCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCATCTCATCTCAACCAAGTTGCTACATAC 1448
1441 TCAGCTACCTACCAAGGCTCTCTGTGTCAACGGGGTAAACATGAGGGAATGCAATTG 1500
1449 TCAGCTACCTACCAAGGCTCTCTGTGTGTCAACGGGGTAAACATGAGGGAATGCAATTG 1508
1501 CAAGAGAGCTGCAATGCTGGAACCAATATGAGGCTTGTGTGATTTGCTGCCACAGAGA 1560
1509 CAAGAGAGCTGCAATGCTGGAACCAATATGAGGCTTGTGTGATTTGCTGCCACAGAGA 1568
1561 AATAATGATTTCTGAGATACAGATGACATGGACCGAAGATATATATGTTGGGGTAG 1620
1569 AATAATGATTTCTGAGATACAGATGACATGGACCGAAGATATATATGTTGGGGTAG 1628
1621 CTTAGCTGAATGTTTATGTTAATTAATATGTTGTTAATTTTGAAGTGAATTTGAA 1680
1629 CTTAGCTGAATGTTTATGTTAATTAATATGTTGTTTCTTAAATTTTGAAGTGAATTTGAA 1688

QY 1681 TGCATCAGCTTCATTAATGCTTTAGAGCGGCGCATATTCGTTTACTAGAACATGAATG 1740
DB 1689 TGCATCAGCTTCATTAATGCTTTAGAGCGGCGCATATTCGTTTACTAGAACATGAATG 1748
QY 1741 AATGTAGTAAATTTTGTGT 1760
DB 1749 AATGTAGTAAATTTTGTGT 1768

RESULT 3
AY038802 1739 bp mRNA linear PIN 16-JUL-2001
LOCUS Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.
DEFINITION AY038802
ACCESSION AY038802
VERSION AY038802.1 GI:14764465
KEYWORDS
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 1739)
AUTHORS Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.
TITLE Biochemical and Molecular Characterization of a Mutation that
Confers a Decreased Raffinose and Phytic Acid Phenotype on
Soybean Seeds
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1739)
AUTHORS Carlson,T.J. and Hitz,W.D.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2001) Crop Genetics, Dupont Co., P.O.Box 80402,
Wilmington, DE 19880-0402, USA

FEATURES
source location/Qualifiers
1..1739
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Wye"
/db_xref="taxon:3847"
/dev_stage="developing embryo"
11..1543
/EC_number="5.5.1.4"
/function="cyclizes glucose 6-phosphate to
L-myo-inositol-1-phosphate"
/note="mIPS-1A"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/translation="MEIENFKYECPNVKYETEIOSVYVYETTELHENRNGTYOMIV
KPKSVKYEKNIHVPLKGVMLVGMGNNGSTLTGTVLANRGIEMATDKIOANVF
GSLTOSAIRVRSFOGEERYAPFKSLIPRVNPDIVFGMDISNNLADAMRAVPD
IDLQOLRPMESMLPLPEITDPDTAAQBERANNVIGTQBOVOOIIKDKAFKE
ARKVKNVVLWANTERYSNLVGLNDTWENLLAVDRNEAISSTLYALAQVENS
PFINSPQTFVPGILDLAIARNTLIGGDFSGQTKMSVLDVPLVAGVIGTIVS
YVHLNNDKMLNSAPOTFRSKISKNVVDWNSNALIYEBGEHPDHVAVIKYVP
GSKRAMDYTSIEFMGKSTVLNHTCDSILAAPIILDIVLIELSTRIEFKANE
GKFSHFPAITLITLITKAPLVPBGTPVNALSKORAMENIMRACVGLAPENMILE
YK"

CDS
11..1543
/EC_number="5.5.1.4"
/function="cyclizes glucose 6-phosphate to
L-myo-inositol-1-phosphate"
/note="mIPS-1A"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/translation="MEIENFKYECPNVKYETEIOSVYVYETTELHENRNGTYOMIV
KPKSVKYEKNIHVPLKGVMLVGMGNNGSTLTGTVLANRGIEMATDKIOANVF
GSLTOSAIRVRSFOGEERYAPFKSLIPRVNPDIVFGMDISNNLADAMRAVPD
IDLQOLRPMESMLPLPEITDPDTAAQBERANNVIGTQBOVOOIIKDKAFKE
ARKVKNVVLWANTERYSNLVGLNDTWENLLAVDRNEAISSTLYALAQVENS
PFINSPQTFVPGILDLAIARNTLIGGDFSGQTKMSVLDVPLVAGVIGTIVS
YVHLNNDKMLNSAPOTFRSKISKNVVDWNSNALIYEBGEHPDHVAVIKYVP
GSKRAMDYTSIEFMGKSTVLNHTCDSILAAPIILDIVLIELSTRIEFKANE
GKFSHFPAITLITLITKAPLVPBGTPVNALSKORAMENIMRACVGLAPENMILE
YK"

ORIGIN
Query Match 97.6%; Score 1717; DB 8; Length 1739;
Best Local Similarity 100.0%; Pred No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 TGGACCTATCAGTGAATTTGTCAAACCCAAATCTGTCAAATACGAATTTAAACCAACAT 223
DB 121 TGGACCTATCAGTGAATTTGTCAAACCCAAATCTGTCAAATACGAATTTAAACCAACAT 180
QY 224 CCATGTTCTTAATTAAGGGGTAAAGCTTGTGGGTGGGAAACAAGGCTCAACCT 283
DB 181 CCATGTTCTTAATTAAGGGGTAAAGCTTGTGGGTGGGAAACAAGGCTCAACCT 240
QY 284 CACCGGTGTATTAATGCTAACCGAGAGGGGATTCATGGGCTACAAAGACAAGATTC 343
DB 241 CACCGGTGTATTAATGCTAACCGAGAGGGGATTCATGGGCTACAAAGACAAGATTC 300
QY 344 ACAAGCCAAATTAATTTGGCTCCCTCACCACCAAGCTTACCTATCCGAGTTGGGCTCTTCA 403
DB 301 ACAAGCCAAATTAATTTGGCTCCCTCACCACCAAGCTTACCTATCCGAGTTGGGCTCTTCA 360
QY 404 GGGAGAGAAATCTATGCCCCATTCAGAACGCTGTTCCAAATGTTAACTCCAGACAT 463
DB 361 GGGAGAGAAATCTATGCCCCATTCAGAACGCTGTTCCAAATGTTAACTCCAGACAT 420
QY 464 TGTGTTGGGGATGGATATCAGCAATGAACTTGCTGATGCGATGCGCAGGGCAAA 523
DB 421 TGTGTTGGGGATGGATATCAGCAATGAACTTGCTGATGCGATGCGCAGGGCAAA 480
QY 524 GGTGTTGACATCGATTTGCAAGAGCAGTTGAGGCTTACATGAAATCATGCTTCACT 583
DB 481 GGTGTTGACATCGATTTGCAAGAGCAGTTGAGGCTTACATGAAATCATGCTTCACT 540
QY 584 CCCCCGAAATCTATGACCCCGGATTTTCATGCTGCCAACCAAGAGCGGTCCAAACAGT 643
DB 541 CCCCCGAAATCTATGACCCCGGATTTTCATGCTGCCAACCAAGAGCGGTCCAAACAGT 600
QY 644 CATCAAGGGCACAAAGACAAGACAAGTCAACAAATCATCAAGACATCAAGGCTTTAA 703
DB 601 CATCAAGGGCACAAAGACAAGACAAGTCAACAAATCATCAAGACATCAAGGCTTTAA 660
QY 704 GGAAGCCACCAAAGTGAACAAGTGTGTACTGTGACTGCCAACACAGAGAGGTACAG 763
DB 661 GGAAGCCACCAAAGTGAACAAGTGTGTACTGTGACTGCCAACACAGAGAGGTACAG 720
QY 764 TAATTTGGTTGGGGCTTAATGACACATGAGAAATCTTGGTGTGTCAGAGAA 823
DB 721 TAATTTGGTTGGGGCTTAATGACACATGAGAAATCTTGGTGTGTCAGAGAA 780
QY 824 TGAGGCTGAGATTTCTCCCTCAACTGTATGCCATTCCTGTGTATGAAAAATGTTCC 883
DB 781 TGAGGCTGAGATTTCTCCCTCAACTGTATGCCATTCCTGTGTATGAAAAATGTTCC 840
QY 884 TTTCTTAATGGAAGCCCTCAGAACATTTGTATACAGGGCTGATGATCTTGGCATGCG 943
DB 841 TTTCTTAATGGAAGCCCTCAGAACATTTGTATACAGGGCTGATGATCTTGGCATGCG 900
QY 944 GAGGAACACTTTGATTTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGT 1003
DB 901 GAGGAACACTTTGATTTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGT 960
QY 1004 GTTGTGTAATTTCTTGTGGGGGCTGTATCAAGGCAACATGTATGCTAGTAAACA 1063
DB 961 GTTGTGTAATTTCTTGTGGGGGCTGTATCAAGGCAACATGTATGCTAGTAAACA 1020
QY 1064 TCTGGGAAAACAATGATGATGAATCTTTCGGCTTCACAAATCTTCCGTTCCAGAAAT 1123
DB 1021 TCTGGGAAAACAATGATGATGAATCTTTCGGCTTCACAAATCTTCCGTTCCAGAAAT 1080
QY 1124 CTCGAAGACCAAGTTGTTGATGATGATGCTCAACAGCAATGCCATCTCTAAGACCTGG 1183
DB 1081 CTCGAAGACCAAGTTGTTGATGATGATGCTCAACAGCAATGCCATCTCTAAGACCTGG 1140
QY 1184 TGAACATCCAGCAATGTTGTTGATTAATTAAGATGCTTACGTAAGGAGGACAGCAAG 1243
DB 1141 TGAACATCCAGCAATGTTGTTGATTAATTAAGATGCTTACGTAAGGAGGACAGCAAG 1200

Oy		1244	AGCCATGGATAGTACACTTCGAGAATATTTCATGTGGTGAAAAGAACACCATTTGTTTGCA	1303
Oy		1304	CAACACATGCAGGAATTCCTCTTTAGCTGCTCCTTAATTCTTGGACCTTGTCCTTCTTGCC	1363
Db		1201	AGCCATGGATAGTACACTTCGAGAATATTTCATGTGGTGAAAAGAACACCATTTGTTTGCA	1260
Db		1261	CACACATGCCAAGAAATTCCTCTTTAGCTGCTCCTTAATTCTTGGACCTTGTCCTTCTTGCC	1320
Oy		1364	TGAGCTCAGCACTAGAAATCGAGTTTAAAGCTGAAAAATGAGGAAAATTCCACTCAATTCGA	1423
Db		1321	TGAGCTCAGCACTAGAAATCGAGTTTAAAGCTGAAAAATGAGGAAAATTCCACTCAATTCGA	1380
Oy		1484	AGTGTGAATGATCATTTGTCAAAGACGCGTGAATGCTGGAAAAACAATAAGAGGCTTGCTG	1543
Db		1441	AGTGTGAATGATCATTTGTCAAAGACGCGTGAATGCTGGAAAAACAATAAGAGGCTTGCTG	1500
Oy		1544	TGGAATTGGCCCCCAGAGAAATAACATGATTTCTCGAGTACAAAGTGAAGCATGGAGCCGAGAA	1603
Db		1501	TGGAATTGGCCCCCAGAGAAATAACATGATTTCTCGAGTACAAAGTGAAGCATGGAGCCGAGAA	1560
Oy		1604	TAAATATAGTTGGGGTAGGCTCTAGCTGAATGTTTTATATGTAATATATATGTTGCTTAATTT	1663
Db		1561	TAAATATAGTTGGGGTAGGCTCTAGCTGAATGTTTTATATGTAATATATATGTTGCTTAATTT	1620
Oy		1664	TTGCAAGTGAATTAAGAAATGCAATGCAAGCTTCATTAATGCTTTAGACGGGGCATATTCTGTT	1723
Db		1621	TTGCAAGTGAATTAAGAAATGCAATGCAAGCTTCATTAATGCTTTAGACGGGGCATATTCTGTT	1680
Oy		1724	TACTAGGAACATGATGATGATGTAGTATTAATTTTGTGT	1760
Db		1681	TACTAGGAACATGATGATGATGTAGTATTAATTTTGTGT	1717
RESULT 4				
LOCUS	BD075269		1533 bp DNA linear	PAT 27-AUG-2002
DEFINITION	Soybean plant producing seeds with reduced levels of raffinose saccharides and phylic acid.			
ACCESSION	BD075269			
VERSION	BD075269.1 GI:22620872			
KEYWORDS	JP 2001519665-A/4.			
SOURCE	unidentified			
ORGANISM	unclassified.			
REFERENCE	1 (bases 1 to 1533)			
AUTHORS	Hitez,W.D. and Sebastian,S.A.			
TITLE	Soybean plant producing seeds with reduced levels of raffinose saccharides and phylic acid			
JOURNAL	Patent: JP 2001519665-A 4 23-OCT-2001;			
COMMENT	EI DU PONT DE NEMOURS AND CO OS Soybean line LR33 PN JP 2001519665-A/4 PD 23-OCT-2001 PF 07-APR-1998 JP 1998543012 PR 08-APR-1997 US 08/835751 PI WILLIAM DEAN HITEZ,SCOTT ANTHONY SEBASTIAN PC G12N15/52,G12N15/82,G12N15/11,G12M9/90,A01H5/00 CC Strandedness: Double; CC Topology: Linear; CC Soybean plant producing seeds with reduced levels of raffinose CC Saccharides CC and phylic acid FH Key Location/Qualifiers FT CDS 1..1533. location/Qualifiers 1..1533 /organism="unidentified" /mol_type="genomic DNA"			
FEATURES				
source				

ORIGIN	/db_xref="taxon:32644"	Query Match	87.0%: Score 1531.4; DB 6; Length 1533;
		Best Local Similarity 99.9%: Pred. No. 0;	
		Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
QY	54	ATGTCATCGAAGATTTTAAAGGTGAGTGTCTTAATGTAAGTACCGAGACTGAGATT	113
Db	1	ATGTCATCGAAGATTTTAAAGGTGAGTGTCTTAATGTAAGTACCGAGACTGAGATT	60
QY	114	CAGTCCGCTGTACAACTACGAAACCAGAACTTGTTCAGGAACAGAAATGGACCTAT	173
Db	61	CAGTCCGCTGTACAACTACGAAACCAGAACTTGTTCAGGAACAGAAATGGACCTAT	120
QY	174	CAGTGGATGTGCAAAACCCAAATCTGTCAATATACGAATTTAAACCACATCATGTTCCT	233
Db	121	CAGTGGATGTGCAAAACCCAAATCTGTCAATATACGAATTTAAACCACATCATGTTCCT	180
QY	234	AAATTAGGGGTAAATGCTTGTGGGTGGGTGGAAACAAACGGCTCAACCTCAACGGTGT	293
Db	181	AAATTAGGGGTAAATGCTTGTGGGTGGGTGGAAACAAACGGCTCAACCTCAACGGTGT	240
QY	234	GTTATTTGCTAACCGAGAGGGCATTTATGCGCTACAAAGACACAGATTCAACACGCAAT	353
Db	241	GTTATTTGCTAACCGAGAGGGCATTTATGCGCTACAAAGACACAGATTCAACACGCAAT	300
QY	354	TACTTTGGCTCCCTCACCAAGCCCTACACTATCCGATTCGAGTGGTCTTCCAGGAGAGGAA	413
Db	301	TACTTTGGCTCCCTCACCAAGCCCTACACTATCCGATTCGAGTGGTCTTCCAGGAGAGGAA	360
QY	414	ATCTATGCCCCCATTCAGAGCGCTGCTTCCAAATGCTTAAACCTTGACGACATTTGTGTTGG	473
Db	351	ATCTATGCCCCCATTCAGAGCGCTGCTTCCAAATGCTTAAACCTTGACGACATTTGTGTTGG	420
QY	474	GGATGGATATCAGCAACATGAACTGTGCTGATGCAAGGACAGGGCAAAAGTGTGTTAC	533
Db	421	GGATGGATATCAGCAACATGAACTGTGCTGATGCAAGGACAGGGCAAAAGTGTGTTAC	480
QY	534	ATCGATTTGGCAAGAGCAAGTGAAGCCCTTAACATGAAATCCATGCTTCCATCCCGGAATC	593
Db	481	ATCGATTTGGCAAGAGCAAGTGAAGCCCTTAACATGAAATCCATGCTTCCATCCCGGAATC	540
QY	594	TATGACCCGGATTTTCAATGCTGTGCCAACCAAGAGAGCGTGCACAAAGTATCAAGAGGC	653
Db	541	TATGACCCGGATTTTCAATGCTGTGCCAACCAAGAGAGCGTGCACAAAGTATCAAGAGGC	600
QY	654	ACAAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGGCGTTTAAAGAAACCAAC	713
Db	601	ACAAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGGCGTTTAAAGAAACCAAC	660
QY	714	AAAGTGGACAAAGTGTGTTACTGTGCACTGCCAACACAGAGAGTACAGTAATTTGGTT	773
Db	661	AAAGTGGACAAAGTGTGTTACTGTGCACTGCCAACACAGAGAGTACAGTAATTTGGTT	720
QY	774	GTGGGCTTTAATGACACCATGAGAAATCTCTTGCTGTGAGACAGAAATGAGGCTGAG	833
Db	721	GTGGGCTTTAATGACACCATGAGAAATCTCTTGCTGTGAGACAGAAATGAGGCTGAG	780
QY	834	ATTTCTCCTTCAACCTGTATATGCAATGCTTGTGTATGAAAAATGTTCTTTTATTAT	893
Db	781	ATTTCTCCTTCAACCTGTATATGCAATGCTTGTGTATGAAAAATGTTCTTTTATTAT	840
QY	894	GGAAAGCCTCAAGAACCTTTTGTACAGGGCTGATTTGATCTTGCCATCGAGAGAACACT	953
Db	841	GGAAAGCCTCAAGAACCTTTTGTACAGGGCTGATTTGATCTTGCCATCGAGAGAACACT	900
QY	954	TTGATTTGGTGAAGTGAATTCACAAAGAGGTGACACCAAAATGAATATCGTGTGGTTGAT	1013
Db	901	TTGATTTGGTGAAGTGAATTCACAAAGAGGTGACACCAAAATGAATATCGTGTGGTTGAT	960
QY	1014	TTCTCTTGGGGGCTGTGATCAACGCAACATCTTATGCTAGTTTCAACATCTTGGGAAC	1073

Db 961 TTCCTTGGGGGGCTGATATCAACCAACATCTATAGTCATTAACCAACCTCTGGAAAC 1020
Oy 1074 AATGATGATGATGATCTTTGGGCTCCACAACTTCCGTTCCAGAAATCTCAAGAGC 1133
Db 1021 AATGATGATGATGATCTTTGGGCTCCACAACTTCCGTTCCAGAAATCTCAAGAGC 1080
Oy 1134 AACGTTGTTGATGATATGCTCAACAGCAATGCCATCTCTATAGACCTGTGTGAATCCA 1193
Db 1081 AACGTTGTTGATGATATGCTCAACAGCAATGCCATCTCTATAGACCTGTGTGAATCCA 1140
Oy 1134 GACCAATGTTGTTGATATGATATGCTCTTACATAGGGGAGACAGCAAGACCAATGAT 1253
Db 1141 GACCAATGTTGTTGATATGATATGCTCTTACATAGGGGAGACAGCAATGACCAATGAT 1200
Oy 1254 GAGTACACTTCAGAGATATTCATGAGTGGTGAAGAGACCACTGTTTGGACACACATGC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGAGTGGTGAAGAGACCACTGTTTGGACACACATGC 1260
Oy 1314 GAGGATTCCTCTTGAAGCTGCTCTATATCTTGAATGCTGCTTCTTGTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTTGAAGCTGCTCTATATCTTGAATGCTGCTTCTTGTGAGCTCAGC 1320
Oy 1374 ACTAGATGAGATTTAAAGCTGAAATGAGGAAATTCATCTATCCACCACTGCT 1433
Db 1321 ACTAGATGAGATTTAAAGCTGAAATGAGGAAATTCATCTATCCACCACTGCT 1380
Oy 1434 ACCATCTCAGCTACCTCCACCAAGGCTCCTGTGTTCCACCGGATACACAGTGTGAT 1493
Db 1381 ACCATCTCAGCTACCTCCACCAAGGCTCCTGTGTTCCACCGGATACACAGTGTGAT 1440
Oy 1494 GCATTGTCAAGCAGCTGCAATGCTGAAACCATATAGAGGCTTGTGTGATTTGGC 1553
Db 1441 GCATTGTCAAGCAGCTGCAATGCTGAAACCATATAGAGGCTTGTGTGATTTGGC 1500
Oy 1554 CCAAGGATTAATATGATTTCTCGATGCAAGTGA 1586
Db 1501 CCAAGGATTAATATGATTTCTCGATGCAAGTGA 1533

RESULT 5
AB032073 1931 bp mRNA linear PIN 08-SEP-1999
LOCUS Nicotiana paniculata NpINPS1 mRNA for myo-inositol-1-phosphate
DEFINITION
AB032073 1931 bp mRNA linear PIN 08-SEP-1999
VERSION
AB032073.1 GI:5834499
KEYWORDS
myo-inositol-1-phosphate synthase.
SOURCE
Nicotiana paniculata
ORGANISM
Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE
1 (bases 1 to 1931)
AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
TITLE Myo-inositol-1-phosphate synthase
JOURNAL Published Only in DataBase (1999)
REFERENCE
2 (bases 1 to 1931)
AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Akiko Hashimoto, Japan Tobacco Inc., Plant
Breeding and Genetics Research Laboratory, 700 Higashibara, Iwata,
Toyoda-cho, Shizuoka 438-0802, Japan
(E-mail: Akiko.Hashimoto@pbgrl.jti.co.jp, Tel: +81-538-32-7116,
Fax: +81-538-33-6046)

FEATURES
source Location/Qualifiers
1..1931
/organism="Nicotiana paniculata"
/mol_type="mRNA"
/db_xref="taxon:62141"
1..1931
/gene="NpINPS1"
92..1624
/gene="NpINPS1"

gene
CDS

ORIGIN
Query Match 63.7%; Score 1120.4; DB 8; Length 1931;
Best Local Similarity 83.0%; Pred. No. 4.2e-291;
Matches 1277; Conservative 0; Mismatches 261; Indels 0; Gaps 0;
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="BA84084.1"
/db_xref="gi:5834500"
/translation="MFLENFVESPNVKYTESHSVYDYOTTELVHDEKXGYOMTV
KPKTKYEFKIDVHPKLGVLVGMGNGNSTLIGVITAREGISMAITKDVQOANVP
GSLTQASTIRVGSFNGSEITAFPKSLPMPVDVVGAGIISPMNLDMARKAFVD
IDLOKOLRPVMSVPLPGIYDPFIANQSSRANVYIKTKKEQIDQIKDIREPK
KNKVDKVVLTWNTATERSNVVYGLNDTMEMLFASVDNREAEISPTLYAJCLENV
PFINSGPONTFVPLIDLAIKNTLIGDDPEKSGOTKMSVLDVFLVAGIKPTISV
YNHIGNDNDGMLISAPOTPRSKEISKSNVNDMSNAILVREGHPHVVYIKVVPV
GDSKRADEYTSSEIFMGSKNTIYVANTCEBSLAAPIILDVTLAEISTRLOKAE
GKFSFHPVATILSYLTKAPLVPGTIPVNLISKRAMENITLACVGLADENNMLE
YK"

Db 51 AATAATGTCATCAGAAATTTTAAAGTTGAGTGTCTTAATGTGAATACCGAGACTGAG 110
Oy 89 AAAATGTTATGAGAAATTTTAAAGTTGAGAGCCCAACGTTAATGACCGAAAGTGA 148
Db 111 ATTCAGTCCGTGTACACTACGAAACCCAGAACTTGTTCAGAGAACAGAAATGCAAC 170
Oy 149 ATTCAGTCCGTGTATGATTTATCAAAACCACTAGTTAGTTATGAGAAATGAGACA 208
Db 171 TATCAGTGTGATTTGCAACCCCAATCTGTCAATATGCAATTTAAACCAATCCATGT 230
Oy 209 TATCAATGAGACCGTCACACCTGACAGTGTCAATATGATGATCAAGCTGATGATAT 268
Db 221 CCTAATATGAGGATATCTGTGAGTTGGGTTGGAACCAACGCTCAACCTCAACCGGT 290
Oy 269 CCCAAATATGAGGATATCTGTGAGTTGGGTTGGAACCAATGATGATCACTGACCGGT 328
Db 291 GGTGTATATGTCACCGAGAGGCACTTCAATGAGGTCACAAAGACACAGATTTCAACAGC 350
Oy 329 GGTGTATATGTCACAGAGAGAAATTTCAATGAGGTCACAAAGATATGATGTCACAGCC 388
Db 351 AATTACTTGGCTCCCTCAACCCAGCCCTCAGCTATCCGATTTGGGTCCTTCAAGGAGAG 410
Oy 389 AATTACTTGGCTCTTACTCAAGGCTTCACTATTCAGATTTGGGTCCTTCAAGGAGAG 448
Db 411 GAAATCTATGCCCCCATTCAGAGCTGCTCCAAATGTTAAACCTTGACAGATTTGTGTT 470
Oy 449 GAGATCTATGCTCCATTTAAAGCTCTCTTCCAAATGTCATTCAGATGAGATGTGTT 508
Db 471 GGGGATGAGATATCAGCAACATGAACTGTGATGCCATGCGCCAGGCAAAAGGTGTT 530
Oy 509 GAGAGATGAGACATCAGCAGATGAAATTTACAGATGCCATGCGCCAGGCTTAAAGGTATTT 568
Db 531 GACATCATTTGACAGACAGCTTGAAGGCTTACATGAAATCCATGCTTCCACTCCCCGGA 590
Oy 569 GATATGTGATCTCAAAACAGCTTGAAGGCTTCAATGAAATCTATGATGCCACTCCGTGT 628
Db 591 ATCTATGACCCGAGATTTGATGCTGCCAACCAAGAGAGCGCTGCCAACAGTCTCAAG 650
Oy 629 ATCTATGACCCGAGATTTGATGCTGCCAACCAAGGATCAGTGCACAAAGTATCAAA 688
Db 651 GGCACAAAGCAAGCAAGTTCACAAATCATCAAGACATCAAGCGCTTTAAGGAGCC 710
Oy 689 GGAACCAAGAAAGAAACAAATGATCAATATCAATTAAGATTAAGGAGATTTAAGGAAAG 748
Db 711 ACCAAAGTGACAGAGTGTGTACTGTGACCTGCCAACACAGAGAGTACATGATTTG 770
Oy 749 AACCAAGTGACAGAGTGTGTACTGTGACCTGCCAACAGAGAGTACATGATTTG 808
Db 771 GTTGGGGCTTAAATGACACCATGAGAGATCTTTGGCTGTGTGACAGAAATGAGCT 830
Oy 809 GTTGTGACCTTAATGACATATGAGAAACCTCTTTGCTTGTGTGACAGAAATGAGCT 868
Db 831 GAGATTTCTCTTCCACCTTGAATGCCATGCTGTGTATGAGAAATGTTCTTTCAAT 890

```
Db      869 GAAATATCTCTCCACTTGTATGCTATTGCTGCATCTTAAATGTCCTTTATT 928
Qy      891 AATGGAAGCCCTCAGAACCTTTTGTACCAAGGCGTATGATCTTGCCATCGGAGAAC 950
Db      929 AATGGAAGCCCCAGAACCTTTGTCTCCAGGCTCATTTATTTGGGCATCAAGAAAG 988
Qy      951 ACTTTATTTGTGGAATGACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTT 1010
Db      989 ACAATTATTTGTGATGACTTTAAAGATGTCAAACCAAAATGAAAGTCAGTCTGTT 1048
Qy      1011 GATTTCTTTGAGGCGGCTGATCAAGCCAACTATATAGTCAGTTACAAACATCTGGA 1070
Db      1049 GATTTCTTTGAGAGCTGATTTAAAGCCAACTATTTGTGACTCAACATTTGGGT 1108
Qy      1071 AACATGATGATGATGAACTTTTGGGCTCCAAACATTTTCGTTCAAGAAATCTCCAG 1130
Db      1109 AACAAATGATGATGATCTGCTGCCCCCTCAAACTTTCCGGTCAAAAGGATCTCGAA 1168
Qy      1131 AGCAAGCTTTGATGATATATGTCACAGCAATGCCATCTCTATAGCTGTGAACT 1190
Db      1169 AGTAATGTTGTGATACATGTTTCAAGCAATGCATCTTTATAGCTTGAAGCAC 1228
Qy      1191 CCAGACCATGTTGTTGTTATTAAGTATGTCCTTAAGTGGGAGACAGCAAGAGCATG 1250
Db      1229 CCGACCACTGTTGTGATTAAGTATGTCCTATATGTGGAGACAGCAAGAGCATG 1288
Qy      1251 GATGACTACACTTCAGAGATATTCATGGGTGAAAGACACATTTTGTGCAACACA 1310
Db      1289 GATGACTACACTCTGAGATTTTCATGGGGGAAAGACCATTTGTTTGGCAATACCT 1348
Qy      1311 TCGAGAGATTTCCCTCTTAAGTGTCTCTATTTATCTTGAAGTTGCTTCTTGTGAGCTC 1370
Db      1349 TCGAGAGATTTCACTTTTAACTGCTCCAAATTAATGATTTGCTTCTTGTGAACTC 1408
Qy      1371 AGCAGTAGATGAGTTTAAAGCTGAAATGAGGAAATTCACACTCATTTCCACCCAGTT 1430
Db      1409 AGTACCCGCACTCAGCTCAAGAGTGAAGAGAGGTAAGTTCCACTCTTCCACCCGTTG 1468
Qy      1431 GCTACCATCTCAGCTACCTCACCAAGGCTCTCTGTGTCACCGGTAACACAGTGTG 1490
Db      1469 GCTACTATCTCAGCTACCTTACCAAGGCTCTCTGTGTAACACAGTACACAGTGTG 1528
Qy      1491 AATGCACTTCAAGCAGCGTGCATGCTGGAAGAAATATAGGGCTTGTGGAATG 1550
Db      1529 AATGCACTTCAAGCAGGAGGGAATGCTTGAGAAATATTTGAGGGCTTGTGACTT 1588
Qy      1551 GCCCAGAGATTAACATGATTTCTCGAGTCAAGTGAAG 1588
Db      1589 GCACGAGAGACACATGATTTCTGGAATACAAATGAAG 1626

RESULT 6
E27176 1950 bp DNA linear PAT 18-JUN-2001
LOCUS E27176
DEFINITION Novel INPS gene derived from Nicotiana.
ACCESSION E27176
VERSION E27176.1 GI:13026394
KEYWORDS JP 199918789-A/1.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 1950)
Shigehiro, Y. and Toshiyuki, K.
Novel INPS gene derived from Nicotiana
Patent: JP 199918789-A 1 13-JUL-1999;
JAPAN TOBACCO INC
OS Nicotiana paniculata
PN JP 199918789-A/1
COMMENT
PD 13-JUL-1999
PF 26-DEC-1997 JP 1997359773
```

```
PR      PR SHIGEHIRO YAMADA, TOSHIYUKI KOMORI
PI      PC C12N15/09, C12N5/10, (C12N15/09, C12R1:91), C12N15/00, C12N5/00,
PC      PC (C12N15/00, C12R1:91)
CC      CC Strandedness: Double;
CC      CC Topology: Linear;
FH      FH Key Location/Qualifiers
FT      CDS 92..1703.
FEATURES
source Location/Qualifiers
1..1950
/organism="Nicotiana paniculata"
/mol_type="genomic DNA"
/db_xref="taxon:62141"
ORIGIN
Query Match 63.7%; Score 1120.4; DB 6; Length 1950;
Best Local Similarity 83.0%; Pred. No. 4.2e-291;
Matches 1277; Conservative 0; Mismatches 261; Indels 0; Gaps 0;
Qy      51 ATAAATGTTATGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGACTGAG 110
Db      89 AAAATGTTTATGAGAAATTTTAAAGTTGAGTGTGAGACCCCAACGTTAAAGTACACCGAAAGTGA 148
Qy      111 ATTCAGTCCGTGTAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 170
Db      149 ATTCAGTCTGTCTATGATTTATCAACCACTGAGTTAGTTCAATGATGATGAAAAATGGACA 208
Qy      171 TATCAGTGAATTTGTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 230
Db      209 TATCAATGAGACCGTCAAGCTTAAGACTGTCAAAATATGATTTCAAGCTGATGATGATGTT 268
Qy      231 CCTAAATTTGGGGTAAATGCTTTGTGGTGGGTGGAACACACGCTCAACCTCACCGGT 290
Db      269 CCCAAATTTGGGGTAAATGCTTTGTGGTGGGTGGAACACAAATGTTCAACCTTGACCGGT 328
Qy      291 GGTGTTATTTGCTAACCGAGAGGGCAATTCATGGGGTCAACAAAGACAAAGATTCACAAAGGC 350
Db      329 GGTGTTATTTGCTAACCGAGAGGAATTTATGAGGCAACAAAGATTAAGGTGCAACAGGC 388
Qy      351 AATTACTTTGGGCTCCCTCAACCAAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAG 410
Db      389 AATTACTTTGGGCTCTCTTACTCAGGCTTACTATTTGAGTTGGGTCTTCAATGAGAA 448
Qy      411 GAAATCTATGCCCATTTCAAGAGCTGCTTCAATGTTTAACTCTGACGACATTTGTGTT 470
Db      449 GAGATCTATGCTCAATTTAAAGCCCTTCCAAATGCTCAATGATGATGATGATGATGTT 508
Qy      471 GGGGGAGGGATATGACGAACATGAACCTGGGTGATGCGATGGCCAGGGCAAAAGGTTT 530
Db      509 GAGAGATGGACATGACGACATGAATTTTGAAGATGCCATGAGCGGCTAAAGTATTT 568
Qy      531 GACATGATTTTGCAGAGAGCTTGAAGCCCTTACATGATTCATGATCTTCCCTCCGGA 590
Db      569 GATATTGATCTCAAAAGCAGTTGAGGCCCTTCAATGATTCATGATTCATGATTCATGATTC 628
Qy      591 ATCTATGACCCGAGTTTCAATGCTGCTGCAACCAAGAGAGCGTCCCAACAGCTCATGAG 650
Db      629 ATCTATGACCCGAGTTTCAATGCTGCTGCAACCAAGAGGTCACGTGCAACCAACGATCAAA 688
Qy      651 GGCACAAAGCAGACCAAGTTCAACAAATCAACAAAGACATCAAGGCGTTTAAAGAGGC 710
Db      689 GGAACCAAAAGAAACCAATTAATCAATCAATTAAGATTAATTAAGGAGGAGGAGGAG 748
Qy      711 ACCAAAGTGAACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 770
Db      749 AACAAAGTGAACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 808
Qy      771 GTTGTGGGCTTAAATGACACATGAGAAATCTTGGCTGCTGTGAGACGAATGAGCT 830
Db      809 GTTGTGAGCTTAAATGACACATGAGAAACCTTTGGCTGCTGTGAGACGAATGAGCT 868
Qy      831 GAGATTTCTCTTCAACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 890
```


Db 869 GAAATATCTCTTCCACTTTGTAGTCTATTGCTGCAATTCCTGAAATGTCCTTTAT 928
QY 891 AATGGAAGCCCTCAGAACACTTTTGTACGAGGCTGATGATCTTGCCATCCGAGAAC 950
Db 929 AATGGAAGCCCTCAGAACACTTTTGTACGAGGCTGATGATCTTGCCATCCGAGAAC 988
QY 951 ACTTTGATTTGGTGGAGATGACTTCAAGAGTGTGACCAAAATGAATCTGTGTGTT 1010
Db 989 ACATTGATTTGGTGGAGATGACTTCAAGAGTGTGACCAAAATGAATGACGTGCTGTT 1048
QY 1011 GATTTCTCTTGGGGGCTGTATCAAGCCAACTATATGTCAATTAACAACATCTGGA 1070
Db 1049 GATTTCTCTTGGGGGCTGTATTAAGCCAACTATATGTCAATTAACAACATTTGGGT 1108
QY 1071 AACAAATGATGATGATGATCTTCTGCGCTCCAAACCTTTCCTGTCAGAGAAATCTCAAG 1130
Db 1109 AACAAATGATGATGATGATCTTCTGCGCTCCAAACCTTTCCTGTCAGAGAAATCTGAAA 1168
QY 1131 AGCAAAGTTTGTGATGATATGTGCAACAGCAATGCCATCTCTATGAGCTGTGAAAT 1190
Db 1169 AGTAATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1228
QY 1191 CCAGACCAAGTTGTTGTTTAAAGTATGTCCTTACGTTAGGAGGACAGCAAGAGAGCATG 1250
Db 1229 CCGACCAAGTTGTTGTTTAAAGTATGTCCTTACGTTAGGAGGACAGCAAGAGGCAATG 1288
QY 1251 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1310
Db 1289 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1348
QY 1311 TCGAGAGATTCCTCTTACCTGCTCTATATCTTGAATGCTGCTTCTTCTGAGCTC 1370
Db 1349 TCGAGAGATTCCTCTTACCTGCTCTATATCTTGAATGCTGCTTCTTCTGAGCTC 1408
QY 1371 AGCACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1430
Db 1409 AGTACCCGATTCAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1468
QY 1431 GGTACCATCTGAGTACCTCAGCAAGGCTCTGCTGTTTCAACCGGAGTACCAAGTGTG 1490
Db 1469 GGTACCATCTGAGTACCTCAGCAAGGCTCTGCTGTTTCAACCGGAGTACCAAGTGTG 1528
QY 1491 AATGATTTGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1550
Db 1529 AATGATCTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1588
QY 1551 GCCCAGAGAAATTAACATGATCTCCGATCAAGTGAAG 1588
Db 1589 GCACCAAGAGAAATTAACATGATCTCCGATCAAGTGAAG 1626

RESULT 7
AB009881 1959 bp mRNA linear PIN 13-JUL-2000
LOCUS AB009881 Nicotiana tabacum mRNA for myo-inositol 1-phosphate synthase,
DEFINITION complete cds.
ACCESSION AB009881
VERSION AB009881.1 GI:8096265
KEYWORDS myo-inositol 1-phosphate synthase.
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum

REFERENCE 1 (sites)
AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
TITLE Screening of wound-responsive genes identifies an immediate-early
expressed gene encoding a highly charged protein in mechanically
wounded tobacco plants
JOURNAL Plant Cell Physiol. 41 (6), 684-691 (2000)
MEDLINE 20399434
PUBMED 10945337
REFERENCE 2 (bases 1 to 1959)

AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1997) Kojiro Hara, Nara Institute of Science and
Technology, Plant Molecular Breeding, Takayama 8916-5, Ikoma, Nara
630-0101, Japan (E-mail:k-hara@bs.ais.tn.ac.jp)
Tel:+81-743-72-5653 (ex. 5653), Fax:+81-743-72-5659)

FEATURES
source
1..1959
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/db_xref="taxon:4097"
/feature_type="leaf"
94..1626
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="BA95788.1"
/db_xref="GI:8096266"

CDS
polyA_site
/note="15 a nucleotides"
1959

ORIGIN
Query Match 63.4% Score 115.6; DB 8; Length 1959;
Best Local Similarity 80.4%; Pred. No. 8.4e-290;
Matches 1307; Conservative 0; Mismatches 319; Indels 0; Gaps 0;

QY 26 CATTATTTCTTAATCTTTGTGAAAAATATGTCATCGAATTTTAAGTTGAGTGTCC 85
Db 66 CTTTATCTTTCTTAAGTTTCTTCAAAATGTTATTAAGAACTTTAAGTTGAAGGCC 125
QY 86 TAAATGAAATGACCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 145
Db 126 CAAGTTAATGACCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 185
QY 146 TGTTCAGAGAAACAGAAATGCACTTATGATGATGATGATGATGATGATGATGATGATG 205
Db 186 AGTTCATGAGGAGAAATGAGCACTTACCAATGAGCTGCAAGCTTAAGCTGCAAAATA 245
QY 206 CGAATTTAAACCAATCCATGTTCTTAATTAAGGTAATGCTTGTGGGTGGGGTGG 265
Db 246 TGAGTTCAAGAGTGAATGTTCAATGTTCCAAATTAAGGGGTAAATGTTGAGTGGGGTGG 305
QY 266 AAACAGGCTCAACCTCACCGGTGTGTTATGCTAACCGAGAGGCAATTCATGGGCG 325
Db 306 AAACAGGCTCAACCTCACCGGTGTGTTATGCTAACCGAGAGGCAATTCATGGGCG 365
QY 326 TACAAAGACAAATTAACCAAGCAATTAATTTGGTCTTCCACCAAGCTTCACTAT 385
Db 366 CACCAAGATTAAGGACCAAGCAATTAATTTGGTCTTCCACCAAGCTTCACTAT 425
QY 386 CCGATTTGGGTCTTCCAGAGAGAGAAATCTATGCCCATTTCAAGACCTGCTTCCAT 445
Db 426 TCGAATTTGGGTCTTCCAGAGAGAGAAATCTATGCCCATTTAAACCTTCTTCCAT 485
QY 446 GGTAAACCTGACGATTTGTTGGGGGATGGGATTCAGCAACATGAACCTGCTGA 505
Db 486 GGTAAACCTGACGATTTGTTGGGGGATGGGATTCAGCAACATGAACCTGCTGA 545
QY 506 TCGCATGGCAGGAGGAGGATTTGACATGATTTGACAGAGAGGATTTGAGGCTTACAT 565
Db 546 TCGCATGGCAGGAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 605
QY 566 GGAATTCATGCTTCCATCTCCCGAATCTATGACCCGATTTCAATGCTGCCAACCAAGA 625


```

Db      606 GGAATCTATGTCCTCCACTACCTGATCTATGATCCGATTCATGCTGCTGAACCAAG 665
Qy      626 GGAAGCGTGGCAACAAGTCTATCAAGGGCAACAAGCAAGCAAGTTCACAAATCATCAA 685
Db      666 GTCAAGCTGGCCAAACGTGATCAAGAAACCAAGAAAGAAACAATGATCAATCATTTAA 725
Qy      686 AGACATCAAGCGCTTTAAGGAAGCCACCAAGGTGACCAAGGTGGTGTACTGTGAGACTGC 745
Db      726 GGAATATTAAGGAGTTTAAGAGAAAGAAACAAGGTGACCAAGGTGGTGTACTGTGAGACTGC 785
Qy      746 CAACACAGAGAGGTACAGTAATTTGGTGTGGGCTTTAATGACACATGAGAGATCTCTT 805
Db      786 TAACACTGAAGATACAGTAATGTGGTGTGTGACTTTAATGATAGTAAATATCTCTT 845
Qy      806 GCGTGTGTGAGACAGAAATGAGGCTGAGATTTCTCTTCCACTTGTATGCCATTCGCTTG 865
Db      846 TGGTCTGTGAGACAGAAATGAGACTGAATATCTCTTCCACTTGTATGCTATTCCTTG 905
Qy      866 TGTATGAGAAATGTTCCCTTTCATTAATGAGAGCCCTGAGAACCTTTGTACAGGGCT 925
Db      906 CATTTGTAATAATGTGCTTTTATTAATGAGAGCCCAACAGAACCTTTGTGCGAGGCT 965
Qy      926 GATTGATCTTGGCCATCGCAGAGAACACTTGTATGTTGTGAGATGACTTCAAGAGTGTCA 985
Db      966 CATGATTTGGCCATCAAGAGAAACAACAGTGTGGTGTGATGACTTTAAGAGTGTCA 1025
Qy      986 GACCAAAATGAATCTGTGTGGTGTGATTTCTTTGGGGGCTGTATCAAGCCAAACATC 1045
Db      1026 AACCAAAATGAAGTCAAGTGTGTGATTTCTTTGGAGCTGTATTAAGCCAAACATC 1085
Qy      1046 TTTAGTCAAGTCAACATCTGTGGAAACAATGATGATGAATCTTTCGGCTCACAAC 1105
Db      1086 AATTGTAGCTACACATTTGGGTAAACAATGATGAATCTTTCGGCTCACAAC 1145
Qy      1106 TTTCCGTTCCAAAGGAATCTCCAAGCAAGCTTGTGTATGATGATGATCAACAGCAATGC 1165
Db      1146 TTTCCGCTCAAAGGAGATCTCGAAAGATATGTTGTATGATGATGATCAACAGCAATGC 1205
Qy      1166 CATCTCTATGAGCTGTGTGAACATCCAGACATGTTGTGTATTAAGTATGTGCTTA 1225
Db      1206 CATCTCTATGAGCTGTGTGAACATCCAGACATGTTGTGTATTAAGTATGTGCTTA 1265
Qy      1226 CGTAAGGGGACAGCAAGAGCCATGATGATGATCACTTCAAGATATTAAGTATGATGATG 1285
Db      1266 TGTGGAGAGACAGAAAGGCAATGATGATGATCACTTCAAGATATTAAGTATGATGATG 1325
Qy      1286 GAGCAACATGTTTGTGACAAACATGCGAGATTCCTCTTGTAGTGTCTCTTATATCTT 1345
Db      1326 GAAACACATGTTTGTGACAAACATGCGAGATTCCTCTTGTAGTGTCTCTTATATCTT 1385
Qy      1346 GGAATTTGCTCTTGTGTGAGCTGAGCTAGCAATGAGTTTAAAGCTGAATAATGAGG 1405
Db      1386 GGAATTTGCTCTTGTGTGAGCTGAGCTAGCAATGAGTTTAAAGCTGAATAATGAGG 1445
Qy      1406 AAAATTTCCACTGATTCACACAGTTGCTACATCTCTGAGTACCTCAACAGGCTCTCT 1465
Db      1446 TAAAGTTCCACTGATTCACACAGTTGCTACATCTCTGAGTACCTCAACAGGCTCTCT 1505
Qy      1466 GGTTCACAGGGGTACACAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1525
Db      1506 GGTACACACAGGTACACAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1565
Qy      1526 CATTAATGAGGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1585
Db      1566 CATTAATGAGGCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1625
Qy      1586 AAGCATGAGGACGAGAAATTAATTAATGATGATGATGATGATGATGATGATGATGATG 1645
Db      1626 AAGATGAGGAGAGACTCAAGAGCTGATGAGGAGAGAAACAATTAATTAATTAATG 1685
Qy      1646 ATATGT 1651
Db      1686 AGCTGT 1691

```

```

RESULT 8
LOCUS      BT013759
DEFINITION Lycopodium esculentum clone 132642F, mRNA sequence.
ACCESSION  BT013759
VERSION     BT013759.1
KEYWORDS   FLI_CDNA.
SOURCE     Lycopodium esculentum (tomato)
ORGANISM   Lycopodium esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopodium.
REFERENCE  1 (bases 1 to 1954)
            Kirkness, E.F., Wang, W. and Vazellie, A.
AUTHORS    Direct Submission
TITLE      Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
JOURNAL    Medical Center Drive, Rockville, MD 20850, USA

FEATURES
            source
                1..1954
                /organism="Lycopodium esculentum"
                /mol_type="mRNA"
                /db_xref="taxon:4081"
                /clone="132642F"
                /issue_type="final"
                /note="TXBY41"

ORIGIN
Query Match      62.1%; Score 1092.2; DB 8; Length 1954;
Best Local Similarity 81.5%; Pred. No. 1.8e-283;
Matches 1265; Conservative 0; Mismatches 288; Indels 0; Gaps 0;

Qy      35 TTAATCTTTGTGAAAATAATATGTCATCGAAGATTTTAAGTGTAGTGTCTTAATGTGA 94
Db      88 TTTGTGTTGTGAGTAATAATGTTTATGAGAACTTTAAGTTGAGAGCCCAATGTTAA 147
Qy      95 GTACACCGAGACTGAGATTCAGTCGCTGTACAACTACGAACCAACCACTTTTACGA 154
Db      148 GTACACGAGAGTGAGATTCAGTCGCTGTATGATTAATCAACCACTGAGTGGTTCATGA 207
Qy      155 GAAACGGAATGACACTATCAGTGTGTGTCAACCCCAATCTGCAATTAAGAAATTTAA 214
Db      208 GGAATAAATGGAACATTCATTAAGAGCTGAAGCCAAACATGTAATAATGAATTTCAA 267
Qy      215 AACCAACATCATGTTCTTAATTAAGGGTAATGCTTGTGGGTGGGTGAAACAACCG 274
Db      268 AACTGATGTCAATGTTCCAAATTAAGAGTTATGCTTGTGATGAGGAGTGAACAATGG 327
Qy      275 CTCAACCTCACCGGTGTATTGCTTAACGAAGGCGCATTTCAATGGGCTTAACAAGGA 334
Db      328 TTCAGCTTGACCGGAGGTATTGCTTAATGAGAGGAATTTCAATGGGCTTAACAAGGA 387
Qy      335 CAAGATTCACAAGCCAAATTAATTTGGCTCCCTCAACCGCTCAGCTATCGAGTTGG 394
Db      388 TAAAGTGACGAAGCCAAATTAATTTGGTCTTTACACAGGCTCTACTATTCGAGTTGG 447
Qy      395 GTCTCTTCAGGAGAGAAATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTAAACC 454
Db      448 GTCTTTCAATGGAAGAGAGATCTATGCTCCCTTTAAGAGCAATCTTCCCATGCTAATGC 507
Qy      455 TACGACATTTGTTGTGGGAGTGGGATATCAAGCAATGAACCTGAGTATGCCATGGC 514
Db      508 AGATGATGATGATTTTGGAGATGGGACATCAAGCAATGAATTTGGAGTGGATGGC 567
Qy      515 CAGGCAAAAGGTTTGAATGATTTGAGAGAGATTTGAGGCTTCAATGGAATTCAT 574
Db      568 AAGATTAAGTTTGTGATGATCTTCAATGCACTGAGGCGGTATGGAATTCAT 627
Qy      575 GCTTCACTCCCGGAATCTATGACCGGATTTCAATGCTGCAACCAAGAGAGCTGTC 634
Db      628 GGTCCACGCTCGGTATCTATGACCTGATTTTCAATGCTGCAACCAAGCTCAAGTGC 687

```

635 CAACACGTCATCAAGGGCACAAGACGACGAGTTCACAAATCATCAAAAGACATCA 694
Db TAAACATGTCATCAAGGAACCAAGAAAGAAACAAATGATCAATATTAAGATATTAAG 747
Qy GGGCTTTAAGGAAGCCACCAAGTGAACAAGTGTGTGATCTGTGACCTGCCAACAACA 754
Db GAGGTTTAAAGGAACCAACAAAGTGAACGAGTGTGTGCTGTGACCTGCCAACAACA 807
Qy GAGGTACAGTAAATTTGTTGGGCTTATGACACATGAGAGAAATCTTTGGCTGTGT 814
Db AAGATACAGCAGT 867
Qy GAGCAAAATAGAGGCTGAGATTTCTCTTCACTTTGATGCAAGGCTGTGTGTGTGTGT 874
Db AATGTACCTTCATCAATGAGAACCCACAAATATCTTTGTCCAGGCTCATTTGATTT 987
Qy TGCATTCGCGAGAAACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 994
Db GGCATCAAGAGAAACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1047
Qy GAAATCTGT 1054
Db GAGTACAGT 1107
Qy TTAACAACATCTGT 1114
Db CTACAACATTTGT 1167
Qy CAAGAAATCTCAAGAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1174
Db AAGAGAGATCTGAAAGAAAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1227
Qy TGAAGCTGT 1234
Db TGAAGCTGT 1287
Qy CAGCAAGAGAGCCATGT 1294
Db TGAAGAGAGAGCCATGT 1347
Qy TGT 1354
Db TGT 1407
Qy CTTTCTGT 1414
Db GCTCTGT 1467
Qy CTGATTTCCACCAAGT 1474
Db CTCTCTGT 1527
Qy GGGTACACCAAGT 1534
Db AGGTACACCAAGT 1587
Qy GGGT 1587
Db GGCTGT 1640

RESULT 9
BT013505 1959 bp mRNA linear PLN 11-MAY-2004
LOCUS Lycopersicon esculentum clone 132193F, mRNA sequence.
DEFINITION BT013505
ACCESSION BT013505.1 GI:47104920
VERSION
KEYWORDS
SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 1959)
AUTHORS Kirkness, E.F., Wang, W. and Vazeille, A.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source location/Qualifiers
1..1959
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132193F"
/tissue_type="fruit"
/note="TZKBF51"

ORIGIN
Query Match 61.8%; Score 1087; DB 8; Length 1959;
Best Local Similarity 80.6%; Pred. No. 4.4e-282;
Matches 1270; Conservative 0; Mismatches 305; Indels 0; Gaps 0;
Qy 36 TAACTTTGTGAAAAATTAATGTTCATCGAGAATTTTAAGTGTGAGTCTTAATGTGAAG 95
Db 59 TGAATTTTTTGTGCAAAATGTTTAAATTTAAAGTGGAAGCCCAATGTGAAG 118
Qy 96 TACACGAGCTGAGATTCAGTCCGTATCAACTAGAAACCAACCACTGTTTCACGAG 155
Db 119 TATATGGAATGAATTCATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 178
Qy 156 AACAGAAATGGACCTATCAGTGAATGTCAAAACCAATCTGTCAATACGAATTTAA 215
Db 179 GAGAGAAATGGAATTCATCAATGATGTGTAAAGCTTAACGTCAATATGAATTTAA 238
Qy 216 ACCAATCTCATGTTCTTAATTAAGGGTAATGCTTGTGGGTGTGGGTGTGAACAACGG 275
Db 239 ACTGATACCCATGTGCCAAAATGGGGGTATGTCTGTGTGAATGGGAGAAACAATGT 298
Qy 276 TCACCTCACCGGTGTGTATTTGCTAACGAGAGGCAATTCATGAGCTTACAAAGAG 335
Db 299 TCACATTTGATCGAGAGGTGTTTTCGAATTCGAAGAAATTTCAATGAGGCAAGAA 358
Qy 336 AAGATTCAACAAAGCCAAATTAATTTGCTCCCTCAACCAAGCTCAGTATCCGAGTTGG 395
Db 359 AAAGTGAACAAGCAATTAATTTTGGTCTCTTACTCAGGCAATCAACCATTCGAGTTGG 418
Qy 366 TCTTTCACGAGAGAAATCTATGCCCATTTCAAGAGCTGCTTCCATGGTTAAACCT 455
Db 419 TCTTTCATATGGGGAAGATCTATGCACTTCGAAAGCCCTTCCTCCATGGTTCAACCA 478
Qy 456 GACGATTTGTGTGTGGGAGTGGGATTCAGCAATGAAACCTGTGATGCCATGAGCC 515
Db 479 GACGATGTATTTTGTGAGATGGGACATTTAGCAATGAAATTTTGGAGATGTATGGTC 538
Qy 516 AGGGCAAGGTGTGTGACATGATTTGCAAGACAGTTGAGGCTTACATGGAATTCATG 575
Db 539 AGGGCTAAAGTTTGTGAAGTTGATCTGCAAAAGCAGTGAAGGCTTACATGGAATTCAT 598
Qy 576 CTTCACATCCCGGAATCTATGACCCGATTTTCAATGCTGTCCCAACCAAGAGAGGTGCC 635
Db 599 GTTCCCTCTCTGTGTATGTAGACCTGACCTTCATTTGGGCTTACCAAGAGAGGTGCC 658
Qy 636 AACAGCTCATCAAGGGCAACAAAGCAAGCAAGTTCACAAATCATCATAGACATCAAG 695
Db 659 AACAGCTGATCAAGGAACCAAGAAAGAAACAGTGTGAACAAATGTTTAAGATTTTAGG 718
Qy 696 GCGTTTAAGAACCAACCAAGTGAACAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 755
Db 719 GAGTTCAAGGAGAAACCAAGTGAACAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 778
Qy 756 AGGTACGTAATTTGTGTGTGGCTTAAATGACACATGAGAGATCTCTGTGCTGTGTG 815

Db		779	AGGTACAGAAATGAGTGGTTGGTCCTTAATGACACCATGGAACCACTTTTACTGCTG	838
Oy		816	GACAGAATGAGGCGTGAGATTCTCTCCTCCACCCTTGATGCCATTGCTGTGTATGGAA	875
Db		839	GATAGAAAATGAGGCTGGAATAATCTCCTCTCACATTGCATGCTAATGCTGTATATGGAA	898
Oy		876	AATGTCTCTTCATTAATGGAAGCCCTCAGAACCTTTTGATCCAGGGCTGATGATCTT	935
Db		889	AATGTGCTTCATCAACGGAAGCCCTCAAAACACTTGTGTCCAGGCTTATTAATTTGG	958
Oy		936	GCCATCGGAGGAAACACTTGTGATTTGGTGGAGATGACTTCAAGAGTGTCCAGCAAAATG	995
Db		959	GCCATTAAGAGAAACACTTAAATTTGTGTGATGACTTTAGAGTGTCTAAACCAAGATG	1011
Oy		996	AAATGTGTTGGTTGATTTCTTGTGGGGGCTGTGATCAAGCCAACATTAATGACT	1055
Db		1019	AAGTCGCGTTGGTTGATTTCTTGTGGAGCCGGATTAAGCCAACGTCATTAATGAGC	1078
Oy		1056	TACAAACATCTGGGAAACAATGATGATGATCTTTGGGCTCCAAAACTTCCGTTCC	1111
Db		1079	TACATCACTTGGGTAACAATGATGAAATCTTCTGCTCCCAAACTTCCGCTCT	1138
Oy		1116	AAGAAATCTCCAAAGAGCAACGTTGTGATGATATGATCAAGAGAAATGCCATCTCAT	1178
Db		1139	AAGAGATCTCAAAGAAATGATATGTTTGTATGACATGTTGCTAGCAAGCCACTTCTAT	1199
Oy		1176	GAGCTGTGTGAACATCCAGACCATGTTGTTGTTAATGATGTCCTTACCTAGAGGAGC	1235
Db		1199	GAATCTGGCGAGCACTTGACCATGTTGTCGATCAAGTATGTTCCATATGTTGAGAGC	1255
Oy		1236	AGCAAGAGAGCCATGATGATGATACCTTCAGAGATTTCAATGGGTGAAAGAGCACTT	1295
Db		1259	AGCAAGAGGGCAATGATGATGATCACTGACAGATTTTCAATGGGCGGAAAGAGCACTATA	1318
Oy		1296	GTTTTGCACAACAATGCGAGGATTCCTCCTTAGCTGCTCCATTAATCTTGACCTGTGC	1355
Db		1319	GTTATGCAACAACCTTGAGGACCTCTCTTTGGAGCTCCAAATATCTTGATTTGGTCTC	1378
Oy		1356	CTTCTGTGAGCTCTCAGCACTAGATCGATTAAAGCTGAAATGAGGGAATTTCCAC	1415
Db		1379	CTTCTCGGTGAATCTAGCACTCGACTTCAGCTCAAAAGCTGAAGGGGAAGGAAATTTCCAC	1438
Oy		1416	TCATTTCCACCAGTTGCTACCATCTCAGCTACCTCACCAGAGCTCTCTGTGTTCCACG	1475
Db		1439	TCTCTCCACCTGTGGCTACTAATCTCAGCTACCTCACAAGAGCTCCCTGGTACACCA	1498
Oy		1476	GGTACACAGTGTGAATGACATTTGCAAAAGAGAGCTGCAATCTGTGGAAAACATATGAG	1535
Db		1499	GGTACTCCAGTGTGAATGCCCTTTCAAGAGAGAGGAGCAATGCTTGAATAATATATGGA	1558
Oy		1536	GCTTGTGTGTGATTGGCCCCAGAGAAATAACATGATTTCTGAGTACAACTGAAAGCATGGGA	1595
Db		1559	GCTTGTGTGTGTTTGGCACCAAGAGAACATGATTTTGGAAATCAATGAAAAATATAT	1618
Oy		1596	CCGAAAGATTAATATA	1610
Db		1619	ATGCAGATGAAGAAA	1633
<hr/>				
RESULT 10				
CPINOIG				
LOCUS	CPINOIG	1978 bp	DNA	linear
DEFINITION	C.paradisel (Macf)			
ACCESSION	Z32632			
VERSION	Z32632.1 GI:602564			
KEYWORDS	INOL gene.			
SOURCE	Citrus x paradiseli			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	Rosidae; eurosoid II; Sapindales; Rutaceae; Citrus.			
REFERENCE	1 (bases 1 to 1978)			
AUTHORS	Holland,D.			

TITLE		Direct Submission
JOURNAL		Submitted (08-APR-1994) Doron Holland, Fruit Tree Breeding and Genetics, Volcani Research, Organisation Center, Bet Dagan, 50250, Israel
FEATURES	location/Qualifiers	
source	1..1978	
	/organism="Citrus x paradisi"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:37656"	
	/clone="c38"	
	/tissue_type="fully expanded leaves"	
	/dev_stage="juvenile plants"	
	188..1711	
gene	/gene="INO1"	
	188..1711	
	/gene="INO1"	
	/codon_start=1	
	/protein_id="CAA83565.1"	
	/db_xref="GI:602565"	
	/db_xref="GOA:P42802"	
	/db_xref="UniProt/Swiss-Prot:P42802"	
	/translation="MPIENFVSES-PNVKDYDTHESHVYDYTELHVNRRGTOMITIKPKVYKPKPDVHVPKLGVLNMGWGNSTLGVYIARBEIGCMATKQVOCANVYFSGSLTQASAIRVGSYNGEBIYAPFKSIIIPKPNPDIVTGGNDISPMNLADAPARVFIIDLOKQLRPWESWVPLPGIYDPFIAAQSRANVYIKGYKKOMQIIOIDIEFKFNKRDVVLMTANTERYSNVYVGLNPTVESLHSLDKNEAIESPSYLAJACLENNPFIDGSPONTVEPGLIDLAIRNCLIGDDEKSGOTKMSVLPFLVQAGIKPISIVPFIKGSNDDNMLSAPORTFSKEISKSNVVDMSNVPFEMGLVNTRPRIKVPYVAAERADDEYTSRIFFMGKSTYIYHNTCESLIAPFIIDVLVAELSTRILQKAEGRKHSFHPVATILSYLTAKPLVPGIVPVNALSQRAMLENILRACVGLPENNMILBYK	
ORIGIN		
Query Match	61.0%; Score 1073.4; DB 8; Length 1978;	
Best Local Similarity	82.0%; Pred. No. 2.1e-278;	
Matches 1275; Conservative	0; Mismatches 271; Indels 9; Gaps 3	
QY	36	TAATCTTTGTGAAAAATPAATGTTTCATCGAGAATTTTAAAGTTGAGTGTCTTAATGTGAAG 95
Db	170	TAAATTTTCTGTGGAAAAATGTTTATCGAAATTTTCAAGGTTGAGAGCCCAATGTGAAG 229
QY	96	TACACCGAGACTGAGATTTCAGTCCGTGTACAACCTACGAAACACCGAATTGTTCAAGAG 155
Db	230	TACACAGATCATGAGATTCACTCTGTGTACATTATGAAACCACTGAGCTTGTTCAAGAG 289
QY	156	AACGAATGCGACCTATACAGTGGATTGTCAAAACCAATCTGCAATAGCAATTTAA 215
Db	290	AACGAAATGCGACTTATACAGTGAATGTCAAGCCCAAACTGTCAAAATAGAGTTCAA 349
QY	216	ACCAACATTCATGTTCCTTAATTAAGGGATATGCTTGTGGTGTGGGTGCAACAACGCGC 275
Db	350	ACGATATTCAATGTCCTTAACCTGGGGTTATGTTGTGGTTGGGTGCAACAACGCGC 409
QY	276	TCAAACCTCACCGGTGTGTTATGCTAACCGAGAGGGCAATTTCATGGCTTACAAAGAGAC 335
Db	410	TCAACTCTCAACGGGTGTGTTATGCGCAATCGAGAGGAAATCTCTGGGCGCAAAAGGAC 469
QY	336	AAGATTCAACAAGGCATTTACTTTGGGTCTCCCTCAACCCAAAGCTAGACTATCCGAGTTGGG 395
Db	470	ACTGTGACGACGACCAATTTATTTGGTTCATTACCCAAAGCATAGAGAAATTAAGATTGGG 529
QY	396	TCTTTCAGGAGAGAGAAATCTATGCCCCATTCAGAGAGCTGTCTTCAATGGTTAAACCT 455
Db	530	TCTTACAAATGAGAGAGATTATATGCTTCATTCAGAGAGCATTCCTCCCTATGGTGAACCA 589
QY	456	GACGACATTTGTTGGGGGATGGGATATGACCAACATGAACTGGCTGATGCCATGGCC 515
Db	590	GATGACATTTGTTGGAGATGGGACATTAAGTACATGAATCTTGCTGATGCCATGGCC 649
QY	516	AGGGCAAAAGGTGTTGACATGCAATTTGCAAGACAGTGTGAGGCGCTTACATGGAATCATG 575
Db	650	AGGGCTAGAGTGTGTTGATTTGATCTTCAAAAGCACTTAAGGCGCTTACATGGAATCATG 709
QY	576	CTTCCACTCCCGGAAATCTATGACCCGGGATTTGATGCTGCCAACCAAGAGAGAGCTGCC 635

```

Db      | 710 GTCCACCTTCGGAATCATATACCCCTGATTTCAATTGCGCAATCAAGGGCGGTGCC 769
Qy      | 636 AACACGCTCATCAGAGGCAACAAAGACAGACAGTTCAACAAATCATCAAAACACTCAAG 695
Db      | 770 AACACGCTCATCAGAGGCAACAAAGACAGACAGTTCAACAAATCATCAAAACACTCAAG 829
Qy      | 696 GCGTTTAAGAGCAACCAAGATGAGCAAGGTGTGTACTGTGAGCTGCCAACAGAG 755
Db      | 830 GAGTTTAAGAGAAAAAAGAGGTGATAGGGTGTGTGTACTGTGAGCTGTCAACCTGAG 889
Qy      | 756 AGGTACAGTAATTTGGTGGGCTTAATGACACACATGAGAAATCTTGGCTGCTGTG 815
Db      | 890 AGGTACAGTAATTTGGTGGGCTTAATGACACATGAGAAAGCCTTCTGCTCACATG 949
Qy      | 816 GACAGAAATGAGGCTGAGATTTCTCTTCCACTCTGTATGCCATGCTGTCTTATGAGA 875
Db      | 950 GACAAAAACGAGCTGAGATTTCTCTTCCACTCTGTATGCCATGCTGTCTTATGAGA 1009
Qy      | 876 AATGTTCTTTCATTAATGAGAGCCCTCAGAACACTTTTGTACAGGGCTGATGATCTT 935
Db      | 1010 AATATCCCTTTTATCATATGAGAGCCCCCAAAACACTTTGTACAGGGCTTATGATCTTG 1069
Qy      | 936 GCCATCGCAGAGAGACCTTTGATGTGAGATGACCTTCAAGAGTGTGACACAAATG 995
Db      | 1070 GGTATTAGAGAGAACGTTTGTATGTTGGTGAATGACCTTCAAGAGTGTGACACAAATG 1129
Qy      | 996 AATCTGTCTGTGTTGATTTCTTGTGGGGGCTGTGATTAACCAACATCTATAGTCACT 1055
Db      | 1130 AATCTGTGTGTGTTGATTTCTTGTGGGGGAGGAAATCAACCAACATCTATGTTAGC 1189
Qy      | 1056 TACAAACATCTGAGAAACAAATGATGATGATCTTTGGCTCCCAAACTTCCGTTCC 1115
Db      | 1190 TACAAACATCTGAGAAACAAATGATGATGATCTTTGGCTCCCAAACTTCCGTTCC 1249
Qy      | 1116 AAGAAATCTCCAGAGCAACGTTGTGATGATGATGATGATGATGATGATGATGATGAT 1175
Db      | 1250 AAGAGATTTCCAAAGTAATGTTGTGATGATGATGATGATGATGATGATGATGATGAT 1308
Qy      | 1176 GAGCTGTGAGAACATCCAGACCAATGTTGTGATGATGATGATGATGATGATGATGATGAT 1235
Db      | 1309 GGGCTGTGAGAACATCCAGACCAATGTTGTGATGATGATGATGATGATGATGATGATGAT 1360
Qy      | 1236 AGCAAGAGAGCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1295
Db      | 1361 ATCGAGAGAGCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1420
Qy      | 1296 GTTTTGCACACACATGCGAGATTCCTCTTACGCTGCTCTAATTAATTTGACCTTGATC 1355
Db      | 1421 GGTCTGCACACACATGCGAGATTCCTCTTACGCTGCTCTAATTAATTTGACCTTGATC 1480
Qy      | 1356 GTTCTGTGAGCTGACGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1415
Db      | 1481 CTGCTGTGAGCTGACGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1540
Qy      | 1416 TCATTCACACACATGCGAGATTCCTCTTACGCTGCTCTAATTAATTTGACCTTGATC 1475
Db      | 1541 TCTTTCACACACATGCGAGATTCCTCTTACGCTGCTCTAATTAATTTGACCTTGATC 1600
Qy      | 1476 GGTACACACATGCGAGATTCCTCTTACGCTGCTCTAATTAATTTGACCTTGATC 1535
Db      | 1601 GGTACACACATGCGAGATTCCTCTTACGCTGCTCTAATTAATTTGACCTTGATC 1660
Qy      | 1536 GCTTGTGTGATGTTGGCTCCCAAGAAATTAATGATTTCTGAGTACAAAGTGAAGCA 1590
Db      | 1661 GCTTGTGTGATGTTGGCTCCCAAGAAATTAATGATTTCTGAGTACAAAGTGAAGCA 1715

```

```

RESULT 11
AF284065 AF284065 1845 bp mRNA linear PLN 30-JUL-2003
LOCUS Sesumum indicum myo-inositol 1-phosphate synthase mRNA, complete
DEFINITION cds.

```

```

ACCESSION AF284065
VERSION AF284065.1 GI:3858815
KEYWORDS
SOURCE Sesumum indicum (sesame)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
ascerids; lamids; Lamiales; Pedaliaceae; Sesamum.
REFERENCE Chun,J.A., Jin,U.H., Lee,J.W., Yi,Y.B., Hyung,N.I., Kang,M.H.,
AUTHORS Pyee,J.H., Suh,M.C., Kang,C.W., Seo,H.Y., Lee,S.W. and Chung,C.H.
TITLE Isolation and characterization of a myo-inositol 1-phosphate
synthase cDNA from developing sesame (Sesumum indicum L.) seeds:
functional and differential expression, and salt-induced
transcription during germination
JOURNAL Planta 216 (5), 874-880 (2003)
MEDLINE 22511750
PUBMED 12624775
REFERENCE 2 (bases 1 to 1845)
AUTHORS Jin,U.-H. and Chung,C.-H.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2000) Division of Biotechnology, Faculty of Life
Science and Resources, Dong-A University, Ha-Dan-Dong, Sa-Gu-Gu,,
Pusan City,
Location/Qualifiers
FEATURES
source
1..1845
/organism="Sesumum indicum"
/mol_type="mRNA"
/db_xref="taxon:4182"
/label_type="Seeds"
/label="authorily: Sesumum indicum L."
51..1583
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAC01148.1"
/db_xref="GI:3858816"
/translation="MFISFVESPENRYTEGEHSYVNYETTELVBESRNGTYOMIV
KPKVYKFEKTDHVPKLGWLVGVGNNNGSTLTCGVANEGISMAKRDVQOANF
GSLVYKFEKTDHVPKLGWLVGVGNNNGSTLTCGVANEGISMAKRDVQOANF
IDLOKOLAPVMEHVPKLGWLVGVGNNNGSTLTCGVANEGISMAKRDVQOANF
ONKDKVYVLTANTERYSNVYVGLNPDIFIANOSRANVYKGRKEQVOOIIDMRPEKE
PFINGSPONTVPGIIDLAIQNSLIGDDPKSOTKXSLVDFIVAGIKPISIVS
YNHLNNDGNNI.SAPQTFRSKEISKSNVDVMSNGILYBGEHPDHI VAKYVYV
GDSKRADEYSEIPIMGKSTVILNHCDESLAPILIDVLAELSTRLOKAGE
GKEHSFHVATILSLYTRAPVPPGT PVNALSQRAMLENILRACVGLADENNMILE
YK"
ORIGIN
Query Match 60.5%; Score 1064.2; DB 8; Length 1845;
Best Local Similarity 80.7%; Pred. No. 6.4e-276;
Matches 1243; Conservative 0; Mismatches 298; Indels 0; Gaps 0;
Qy 47 AAAAATATGTCATCGAGATTTTAAGTTAGTGTCTTAATGTAAGTACACCGAGAC 106
Db 44 AACGAAATGTCATCGAGAGCTTCAAGTTGAGAGCCCAATGTCAGTTACACCGAGAG 103
Qy 107 TGAGATTCAGTCGCGTCAACTACGAAACCAACCGAATGTTTCCAGAGAACGAGATG 166
Db 104 TGAGATTCAGTCGCGTCAACTACGAGACTACGAGAGCTTGTTCATGAGAGAGAAAG 163
Qy 167 CACTATCAGTGTGTCACAAACCAAACTGTCAAAATCGAATTTAAACCAACATCGCA 226
Db 164 CACTATCAGTGTGTCACAAACCAAACTGTCAAAATCGAATTTAAACCAACATCGCA 223
Qy 227 TGTTCCTAAATTAAGGGGTATGCTTGTGGGTGGGAGTGAACCAACGCTCAACCTCAG 286
Db 224 TGTTCCTAAATTAAGGGGTATGCTTGTGGGTGGGAGTGAACCAACGCTCAACCTCAG 283
Qy 287 CGTGTGTATTTGCTAACCGAGAGGCAATTCATGGGCTACAAAGACCAAGATTCACA 346
Db 284 TGGCGGTGCATTTGCAACAGGGAAGGATTTCAATGAGGCAACAAAGATTAAGGTGACA 343
Qy 347 ACCCAATTAATTTGGCTCCCTGACCCAAAGCTCAGCTATCCAGTTGGGTCTTCCAGGG 406

```

Db 344 GGCATTTATTTGGGCTGACGACGAGCATCTCTCAATCAGAGTTGGTCTTCAATGG 403
 Qy 407 AGAGAAATCTATCCCATTTCAAGAGCTGCTTCCATGGTTAACCCTGACACATTTGT 466
 Db 404 AGAAGAGATCTATGCCCCCTTCAAAAGCCCTTCTCCATGATGTAACCCAGATGACGTAGT 463
 Qy 467 GTTTGGGGGATGGGATATGCAACATGAACTGGGCGATGCGATGGCCGAGGCAAAAGT 526
 Db 464 GTTTGGGGGATGGGATATGCAATGAAATTTGGCGATGCGATGGGCGAGGCAAGT 523
 Qy 527 GTTTGACATGATTTGGCAGAGAGAGTTGAGGCTTACATGAAATCCATGCTTCACTCCC 586
 Db 524 GTTAGATATTTGATCTCCAGAGAGAGCTGAGGCCCTACATGAGACACATGTTCCCATCCC 583
 Qy 587 CGGAATCTATGACCCCGGATTTCAATGCTGCGCAACAGAGAGCGTCCCAACAGCTCAT 646
 Db 584 TGGAAATCTAGATCTGATTTCAATAGCCGCCCAATCAGGAGATCACTGTCCAACAGTGAT 643
 Qy 647 CAAGGCGACAAAGCAGAGAGAGATTTCAACAATCATCAAGACATCAAGGCGTTTAAGGA 706
 Db 644 TAAAGAACCCAAAGAAAGAAAGATTTCAACAGATCATCAAGACATGAGGACCTTTAAGGA 703
 Qy 707 AGCCACCAAGTGGACAAAGGTGGTGTGATCTGTGACCTGCGCAACAGAGAGTACAGTAA 766
 Db 704 GCAAAATTAAGTGTGACAAAGTGTGATGCTGTGTGACAGCCACACAGAAAGATACAGTAA 763
 Qy 767 TTTGTTGTGGGCTTAAATGACAACAATGAGAAATCTTTGGCTGTGTGACAGAAATGA 826
 Db 764 TTTGTTGTGGGCTTAAATGACAACAATGAGAAATCTTTGGCTGTGTGACAGAAATGA 823
 Qy 827 GGTGATGATTTCTCTTCCATCTGTATGTCATGCTTGTGTATGAGAAATGTTCTT 886
 Db 824 GGGCGAATATCTCTTCAACCTTGAATGCAATGCTGTGTGTGAAATGTTCCGTT 883
 Qy 887 CATTAATGGAAGCCCTCAGAACATTTTGTACAGAGGCTATGATCTTCCATCCGAG 946
 Db 884 CATTAATGGAAGCCCTCAGAACATTTTGTGTGAGAGGCTATGATCTTCCATCCGAG 943
 Qy 947 GAACACTTGTATGTGAGAGATGACTTCAAGATGCTGACAGCCAAATGAAATCTGTGT 1006
 Db 944 GAACACTTGTATGTGAGAGATGACTTCAAGATGCTGACAGCCAAATGAAATCTGTGT 1003
 Qy 1007 GGTGATTTCTCTTGGGGGCTGTATCAAGCCACATCTATGTCATGTTACACCATCT 1066
 Db 1004 AGTTGATTTCTTGTGTGAGCTGTGTATCAAGCCACATCTATGTCATGTTACATCT 1063
 Qy 1067 GGGAAACAATGATGTATGATCTTTGCGGCTCACAACATTTTCCGTTCCAAAGAAATCTC 1126
 Db 1064 GGGAAACAATGATGTATGATCTTTGCGGCTCACAACATTTTCCGTTCCAAAGAAATCTC 1123
 Qy 1127 CAAGAGCAACGTTGTGATGATGATGTCACAGCAATGCCATCTCTATGAGCGTGTGA 1186
 Db 1124 TAAAGCAATGTGTGATGATGATGTCACAGCAATGCCATCTCTATGAGCGTGTGA 1183
 Qy 1187 ACATCCAGACATGTTGTTGTTATTAAGTATGTCCTTACGTAGGAGACAGCAAGAGC 1246
 Db 1184 ACACCCCGACATATGTTGTCTATCAAGTATGATCGTATGTTGGGGGACAGCAAGAGC 1243
 Qy 1247 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1306
 Db 1244 AATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1303
 Qy 1307 CACATGCGAGATTTCCCTTTAGCTGCTCTATATCTTGGACCTTGTCTTCTTCTGA 1366
 Db 1304 CACGTGTGAGGATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1363
 Qy 1367 GCTCAGACATGAGATGAGTTTAAAGCTGAAAATGAGGAAAATTCATCTCATTCACCC 1426
 Db 1364 ACTCAGACATCTGATTCAGCTCAAGCCGAGGAGAGGAGGCAAAATTCATCTGTTCCACC 1423
 Qy 1427 AGTTGTACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1486

Db 1424 GGTGCAACATCTAGCTACTTAACCAAGCCCTTGTGATCTCAGAGCACCGGT 1483
 Qy 1487 GGTGATGCAATTTGTCAAGACAGGCTGATGCTGAGAAAATCATATGAGGCTTGTGTGG 1546
 Db 1484 CGTCAACGACATCTTACAGAGGCTGATGCTGAGAAATCTTGAAGGCTTGTGTGG 1543
 Qy 1547 ATTAGCCCGCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1587
 Db 1544 ATTAGGCTCCAGAAACATATGATTTTGGAAATACAGTGA 1584
 RESULT 12
 MCUS2511
 LOCUS
 DEFINITION
 Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase
 mRNA, complete cds.
 ACCESSION
 VERSION
 U32511
 U32511.1 GI:975887
 KEYWORDS
 SOURCE
 ORGANISM
 Mesembryanthemum crystallinum (common iceplant)
 Mesembryanthemum crystallinum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Astaceae; Mesembryanthemum.
 REFERENCE
 AUTHORS
 1 (bases 1 to 2053)
 Ishitani, M., Majumder, A. L., Bornhouser, A., Michalowski, C. B.,
 Jensen, R. G., and Bohner, H. J.
 Coordinated transcriptional induction of myo-inositol metabolism
 during environmental stress
 Plant J. 9 (4), 537-548 (1996)
 JOURNAL
 MEDLINE
 PUBMED
 96208959
 8624516
 2 (bases 1 to 2053)
 REFERENCE
 AUTHORS
 Ishitani, M., Majumder, A. L., Bornhouser, A., Michalowski, C. B.,
 Jensen, R. G., and Bohner, H. J.
 Direct Substitution
 Submitted (27-JUN-1995) Hans J. Bohner, Biochemistry, University
 of Arizona, Biosciences West 516, Tucson, AZ 85721, USA
 TITLE
 JOURNAL
 FEATURES
 Location/Qualifiers
 1..2053
 /organism="Mesembryanthemum crystallinum"
 /mol_type="mRNA"
 /db_xref="taxon:3544"
 182..1720
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase"
 /protein_id="AAB03687.1"
 /db_xref="GI:975888"
 /translation="MTESFKYSPVRYTENEISVYVDTTELHNRKQAGYQW
 IYKPTVQYHETDTRVPLGVMVGWGNSTLIGVIAIRBGIWATKDKIQAN
 YFGSLTQASISIVGSPNGEIVAPFKSLPMWNPDPVVGWDIDSMNLADMTARV
 PDIDLOKQLRPYMEHNVPLPGIYDDPIAANGSANNVYIKETKEOVERYKIDREF
 KEKNKYDQVNVNTERYSNVYVGLDGMENLNLASLEKSEISPSVYALCIES
 NIPFNGSPQNTFVPEGLDIAIKKSLGDDPFKSGQTKSVLVDPLVAGIKPYSI
 VSNVHLGNDDGNLSAPQTFRSEIKSVYVDWYASNIIYEPREHPIVYIKVP
 YVGDKRANDERTSSIEMGNTIVWANTCSLLAAPILDLVLAELSTIIQLAE
 EEDKHSFHPVATILSYLTKAPLVDPGPVVAALSKORAMLENILRACVGLAPENNMI
 LEYK"
 ORIGIN
 Query Match 59.0%; Score 1037.8; DB 8; Length 2053;
 Best Local Similarity 77.6%; Pred. No. 8.7e-269;
 Matches 1282; Conservative 0; Mismatches 362; Indels 7; Gaps 2;
 Qy 44 GTGAAAATTAATGTCATGAGAAATTTAAGTTGAGTGTCTCTTAATGTAAGTACACGA 103
 Db 172 GAGAGAGAAATGTTTATGAGAGCTTCAAAATGAGAGCCAAAGCTGAATGACAGAA 231
 Qy 104 GACTGAGATTGAGTGTGATCACTGAAACCAACCGAATCTTGTTCAGAGAACAGAA 163
 Db 232 GATGAGATTGATGAGTGTACACTGACACCAACCGAGTTGGTTTCATGAGATGCAA 291
 Qy 164 -----TGGCACTATCACTGATGTGTCAACCAACCTGTCAAAATGCAATTTAAAC 217

```
Db 292 AGATGCTGGTGTATCAATGATCGTTAAGCCCAAGACCGTTAGATCACTTCAGAC 351
Qy 218 CAACATCATGTTCTTAATTAGGGTAAATGCTTGGGTTGGGGTGAACAAGGGCTC 277
Db 352 TGACACTGCTGCCCCAACTCGGGGTATGCTGTGTGATGGGGTGAATAACGGCTC 411
Qy 278 AACCCCTACCGGTGGTGTATGCTTAACCGAGAGGGCATTTTCATGGGCTCAAAAGACA 337
Db 412 CACCTTCACAGGTGTGTGATGCGCAACCGAGAGGAATCTTGGGCAACCAAGACA 471
Qy 338 GATTCAACAGCCAAATTACTTTGGCTCCTCAACCGACTCACTATCCAGTTGGGTG 397
Db 472 GATCAGCAAGCACTATTTGTGTGCTGTGACCCAGGCTTCTCAATTCGGGTGGGTC 531
Qy 398 CTTCAGAGAGAGAAATCTATGCCCATTCAGAGCGCTTCCTCAATGGTTAACCTGA 457
Db 532 TTTTAATGAGAGAGATTATGCTCTTTCAGAGCTGTCTCCCTATGTGAACCCAGA 591
Qy 458 CGACATGTGTTGGGGATGGAATTCAGCAACATGAACCTGCTGATGCCATGGCCAG 517
Db 592 TACAGTAGTGTGGGGTGGGACATTAAGCAGATGAACCTGCTGATGCATGACACAG 651
Qy 518 GGCAGAGGTGTTGATCATGATTTGCAAGAGAGTGAAGGCTTACATGGAATCCATGCT 577
Db 652 GGCTAGGGCTTTCATGATGATCTTCAAAAGCAGCTCAGGCTTACATGAGACATAGGT 711
Qy 578 TCCACTCCCCGGAATCTATGACCCCGATTTCAATGCTGCCAACCAAGAGACGTCGCA 637
Db 712 TCCCTCCCTGTGATTTAGACCCCTGATTTCAATGCTGCGAACCAAGGCTCCCGACCA 771
Qy 638 CAACGTCATCAAGGGGCAAGAGAGCAAGTTCACAAATCATCAAGACATCAAGGC 697
Db 772 CAACGTCATTAAGGGGCAACAGAGAAACAAGTTGAGCAGTCATTAAGACATTAAGGA 831
Qy 698 GTTTAAGAGAGCCACCAAGTGAACAAGTGTGTTACTGTGATGCTGCCAACAGAGAG 757
Db 832 GTTCAAGAGAGAACCAAGTGAACAAGTGTGCTATGAGATGCGCAACAGAGAG 891
Qy 758 GATACGATATTGTTGTGTGGGCTTAATGACACCATGAGGAATCTTGGCTGCTGTGA 817
Db 892 GATACGCAATGTGTGTGGGCTCAACAGACACATGAGGAACCTGTGTGCTCTGTGA 951
Qy 818 CAGAAATGAGGCTGAGATTTCTCTTCCACTGTATGTCATGTTGTTGTTATGAGAAA 877
Db 952 GAAAAAGAGTGGAGATTTCTTCATCCTGCTGTATGCTTGGCATGTATCGAAGAAA 1011
Qy 878 TGTTCCTTTCATTAATGAGAGCCCTCAGAACATTTTGTACAGGGCTGATGTATCTGC 937
Db 1012 TATTCCTTCATCAATGTGATGAGCCCTCAGAACACCTTGTGCGAGGGCTCATGTATCTGC 1071
Qy 938 CATCGAGAGAACCTTTGATGTGTGAGATGATCTTCAAGAGTGTGACAGCAAGATGAA 997
Db 1072 TATCAAGAGAAACAGTTGATGTGTGTGATGATCTTCAAGAGTGTGACAGCAAGATGAA 1131
Qy 998 ATCTGATGTGTTGATTTCTTGTGGGGCTGATCAAGCCAACTCTATAGTCACTTA 1057
Db 1132 ATCTGCTCTGTGATTTCTCTGTGGGGCTGATATTAAGCCAACTCTATAGTCACTTA 1191
Qy 1058 CAACCATCTGGGAACAATGATGATGATCTTTCGCTCACAACCTTTCGTTCCA 1117
Db 1192 CAACCATCTGGGAACAATGATGATGATGATCTTTCGCTCACAACCTTTCGTTCCA 1251
Qy 1118 GGAATCTCCAGAGCAACGTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1177
Db 1252 GGAGATTTCTAAGAGCAACGTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1311
Qy 1178 GCTGTGTGAACATCCAGACATGTTGTGTTATTAAGTATGATGATGATGATGATGATGATGAT 1237
Db 1312 GCGGGGCAACACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1371
Qy 1238 CAAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1297
```

```
Db 1372 CAAGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1431
Qy 1298 TTTGACAAACATGCGAGAGATTCCTCTTATGCTGCTCTTATTAATCTTGAATGATGATGATGAT 1357
Db 1432 CATGACAAACCTGTGAGGATCTCTCTTGGCTGCTCCCATCATCTGATGATGATGATGATGATGAT 1491
Qy 1358 TCTTCTGAGCTGACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1417
Db 1492 CTTGCTGAACATGACACCGGATTCAGCTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1551
Qy 1418 ATTCACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1477
Db 1552 CTTCATCGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1611
Qy 1478 TACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1537
Db 1612 CACTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1671
Qy 1538 TTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1597
Db 1672 TTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1731
Qy 1598 GAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1656
Db 1732 CAAGAGCCACACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1791
Qy 1657 TATTAATTTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1687
Db 1792 TGTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1822

RESULT 13
AY028259 1845 bp mRNA linear PLN 26-MAR-2001
LOCUS Avicennia marina myo-inositol 1-phosphate synthase mRNA, complete
DEFINITION cds.
ACCESSION AY028259
VERSION AY028259.1 GI:13447456
KEYWORDS
SOURCE
ORGANISM
Avicennia marina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Acanthaceae; Acanthaceae incertae
sedis; Avicennia.
REFERENCE
1 (bases 1 to 1845)
Jithesh,M.N., Parani,M. and Parida,A.
Characterization of a cDNA for myo-inositol 1-phosphate synthase
from the mangrove species Avicennia marina
Unpublished
2 (bases 1 to 1845)
Jithesh,M.N., Parani,M. and Parida,A.
Direct Submission
Submitted (02-MAR-2001) Plant Molecular Biology Lab, M.S.
Swaminathan Research Foundation, III Cross Street, Chennai 600 113,
India

FEATURES
source
1..1845
location/Qualifiers
/organism="Avicennia marina"
/mol_type="mRNA"
/db_xref="taxon:82927"
27..1556
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAK21969.1"
/translation="MTESFKVDSRVKXTDEISHSVNYETAEELVHENRNGTYQNIY
KRTKYVERTDTHVYKLGVMVGGNNGSTLTGTVLANRGIISWATDXTQOANYF
GSLTDSATIRVGSFNGBEIYAPFKSLPVPVNDIIFGSDISNNMLADAMGOGELD
IDLOQRPRYMHMPLRGIVYDPDLTANOGSRANNVTKTEKBOVOQVYIKDKERE
KTKERVVGTATNTERYSNVVGLDPTAESLASIERNEAIESPSTLYAIACVDEVP
FNGSPONTFVGLIELAIQRNCLAGDFFKSGGOTKMSVLVDFVAGIKFTISYVS
NHIGNDGNILSAPOTFRSKEISKSNVDDWVANSILYEPBHPDHVVIVYDPVG
```


ORIGIN

DSKRAMEYSEIFMGKSTIYVHNTCEDSLAIPILDLVLLAEHSTRIQLKAESEG
KPHSFHVAITILSYLTKAPLVPPTVNVNALSQRAMLENILIRACVGLPENNMILEY
K"

Query Match 58.9%; Score 1037.4; DB 8; Length 1845;
Best Local Similarity 80.0%; Pred. No. 1.1e-268;
Matches 1234; Conservative 0; Mismatches 306; Indels 3; Gaps 1;

QY 44 GTGAAAAAATATATGTATGAGAAATTTTAAGTTGAGTCTTAATGTGAATGACCGA 103
DB 17 GAGAGAAATATGTTATCGAGAGCTTCAAGGTTGACAGCCCAAGCTTAAGTACAG 76
QY 104 GACTGATTCAGTCCGTGTACAACTACGAAACACGAACTTGTTACAGAGAACGAA 163
DB 77 CAGTGAATTCACCTCGTCTACACATATGAAATCTGACAGACTGTGTACAGAACGAA 136
QY 164 TGGACCTTATCAGTGTATGTCAAAACCAATCTGTCAATATGAAATTTAAACCAACAT 223
DB 137 CGGTACTATCAGTGTATGTCAAAACCAATCTGTCAATATGAAATTTAAACCAACAT 196
QY 224 CCATGTTCTTAATTAAGGGGTATGCTGTGGGTGGGTGAAACAAGGCTCAACCT 283
DB 197 CCATGTCCTGAATTAAGGTATGCTGTGGGTGGGTGAAACAAGGCTCAACCT 256
QY 284 CACCGGTGTATTTATGTTAACCGAGAGGGCAATTCATGGGCTTCAAGAGCAAGTTCA 343
DB 257 GACAGGTGTATTTATGTTAACCGAGAGGGCAATTCATGGGCTTCAAGAGCAAGTTCA 316
QY 344 ACAAGCAATTAATTTGCTCCCTCAACCAAGCTTCAAGTATCGAGTTGGTCTTCCA 403
DB 317 ACAGGCAATTAATTTGCTCCCTCAACCAAGCTTCAAGTATCGAGTTGGTCTTCCA 376
QY 404 GGGAGAGAAATCTATGCCCATTCAGAGCTGCTTCAATGTTTAACTTGAACACAT 463
DB 377 TGGAGAGAGATCTATGCTCAATCAAGAGCTTCAATGTTTAACTTGAACACAT 436
QY 464 TGGTGTGGGGATGGGATATACGAAACATGAACTGGGTGATGCCATGGCCAGGCAAA 523
DB 437 AATTTTGGGGATGGGATATACGAAACATGAACTGGGTGATGCCATGGCCAGGCAAA 496
QY 524 GGTGTTTGCATGATTTGCAGAGAGAGTGAAGCTTATCATGGAATCCATGCTTCCACT 583
DB 497 AAGGTATGACATGATCTCCAAAAGAGCTGAGGCCCTTACATGAGACATGATCCCACT 556
QY 584 CCCCAGATCTATGACCCGGAATTCATGCTGCAACCAAGAGAGCGTCCAAACAGCT 643
DB 557 CCTGGGATCTACTACCTGACTTCATGCTGCAATCAGGGCTCAAGTCCAAACAGCT 616
QY 644 CATCAAGGCAACAAAGCAAGCAAGTTCACAAATCATCAAAAGATCAAGGCGTTTAA 703
DB 617 GATCAAAAGAACCAAGAGAGAGAGTTCACAAAGTATCAAAAGATTAAGAGTTTAA 676
QY 704 GGAAGCAACCAAGTGAAGAGTGTGTACTGTGTAAGTCCCAACACAGAGAGTACAG 763
DB 677 GGAAGAAATTAAGTGAAGAGTGTGTAAGTCCCAACACAGAGAGTACAG 733
QY 764 TAAATTTGTTGGGCTTAAATGACACATGAGAAATCTTGGCTGTGTGACAGAA 823
DB 734 CATGTGTTGGGCTTAAATGATACAGCGGAAACCTTATGCGCTGTGGAAGAA 793
QY 824 TGAAGGTGATTTTCCCTTCCACCTGTATGATGCTTGTGTATGAAAAATGTTCC 883
DB 794 TGAAGGTGATTTTCCCTTCCACCTGTATGATGCTTGTGTATGAAAAATGTTCC 853
QY 884 TTTCATTAATGAGAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGCCATCCG 943
DB 854 TTTTATCAATGACAGCCCAAAACATTTGTTCCAGGCTGATGATGATGATGATGATGAT 913
QY 944 GAGGAACACTTTGATTTGTGAGATGATCTTCAAGAGTGTCAAGACCAATGAAATCTGT 1003
DB 914 AAGGAACCTTTGATTTGTGAGATGATCTTCAAGAGTGTCAAGACCAATGAAATCTGT 973

QY 1004 GTTGTGATTTCTTGTGGGGCTGTATCAAGCAACATCTATAGTACAGCA 1063
DB 974 GCTAGTTGATTTCTTGTGGGGCTGTATCAAGCAACATCTATAGTACAGCA 1033
QY 1064 TCTGGGAAACATGATGATGATTTTGGCTTCCAAACATTTCCGTTCCAGGAAAT 1123
DB 1034 TTTGGGAAATATGATGATGATTTATCTACAGCCCGCAACATTTCCGGTCAAGAGAT 1093
QY 1124 CTCCAGAGCAACGTTGTTGATGATTTGCTCAACAGCAATGCCATCTCTATGAGCTTG 1183
DB 1094 CTCCAAACCAATGTTGATGATGATTTGCTCAACAGCAATGCCATCTCTATGAGCTTG 1153
QY 1184 TGAACATCCAGACATGTTGTTGATTTAAGTATGATGATGATGATGATGATGATGAT 1243
DB 1154 AGAGACCTGTATCTATGTTGTTGATTTCAAGTATGATGATGATGATGATGATGATGAT 1213
QY 1244 AGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1303
DB 1214 AGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1273
QY 1304 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1363
DB 1274 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1333
QY 1364 TGAAGTCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1423
DB 1334 GGAATCTACAGACCCGATTCAGTCAAGGCTGCTGTTCCACCGGATCAAC 1393
QY 1424 CCCAGTGTACATCTTCCAGTCAAGGCTGCTGTTCCACCGGATCAAC 1483
DB 1394 TCCCGTGGCACATCTCAGTCAAGGCTGCTGTTCCACCGGATCAAC 1453
QY 1484 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1543
DB 1454 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1513
QY 1544 TGAATTTGGCCGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586
DB 1514 TGAATTTGGCCGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556

RESULT 14
AF433879
LOCUS
DEFINITION
Suaeda maritima subsp. salsa myo-inositol-1-phosphate synthase
(INPS) mRNA, complete cds.
ACCESSION
AF433879
VERSION
KEYWORDS
SOURCE
ORGANISM
Suaeda salsa
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Suaeda.
REFERENCE
1 (bases 1 to 1845)
Wang, J.P., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.
TITLE
Direct Submission
JOURNAL
Submitted (15-Oct-2001) The Biology Department of Shandong Normal
University, Key Laboratory of Plant Stress Research, No.88, Wenhua
East Road, Jinan, Shandong Province 250014, China

FEATURES
source
1..1986
/organism="Suaeda salsa"
/mol_type="mRNA"
/sub_species="salsa"
/db_xref="taxon:126914"
1..1986
/gene="INPS"
/gene="INPS"
140..1672
/gene="INPS"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAU28131.1"


```

/db xref="GI:16755889"
/translation="MFIENFKVESPNVYNTENEIHSVYVYETTELVHENRKGSGYOMI
VPEKTVKQFOKTOHPKLGVMVVGKNGNCTLTGVTGVTANEGSMATKDKIOOANY
RGLTQASIRVSGNGBEITVAPKSLIPNVDDIVREGMISPLINLADMTARAKVL
DIDLOKQAPRYEHWNPUBGIYDPPIIANDNSRANNIITKTKRQOEVOYIKDIEFK
ENKVDKVVLTANTATERYSDVVGILNTTENLMAVENESISPTLYATACIOED
VPIFGSPQNTFVPLI,IELAKNCLIGDDPKSGQVPLVPLFVPLVPLVPLVPLV
SYNHLGNDGNML,SAPOTERKSEIKSNVDDMVASNGILYEPGEHPHVVVYKVPY
VDSKRAMDEYSEIFMGKNTIVLNTCEDSLAAPIILDLVLAEI,STRIEFAEET
DKFHFHFHVATILISYLTKAPLVPPTPVNALSKQRMLEKRAOCVGLAPENNMLE
YK"

```

ORIGIN

```

Query Match      58.5%; Score 1030; DB 8; Length 1986;
Best Local Similarity 80.2%; Pred. No. 1.1e-266;
Matches 1236; Conservative 0; Mismatches 300; Indels 6; Gaps 2;

```

```

QY 50 AATPATGTCATCGAGAAATTTTAAGTTAGTGTCTTAATGTGAAGTACACGAGACTGA 109
DB 136 AAAAAATGTCATCGAAAAATTTCAAGGTAGAAAAGCCCAATGTTAACTACAGAAATGA 195
QY 110 GATTCAGTCGGTGCACATCGAAACCAACCAACTGTTCACGAGAACGGAATGGCAC 169
DB 196 AATTCATTAGTTACATTAATGAAACCACTGAATTAATTCAGAGAAATCGAAAGCTC 255
QY 170 ---CTATCAGTGGATTTGTCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCA 226
DB 256 TGGTATCATAGATGATCGTCAAACTGAAACCGTCAATTCAGAACTCAAACTCA 315
QY 227 TGTTCCTAAATTAAGGGTATGCTGTGTGGGTGGGTGGAACCAACGCTCAACCTCAC 286
DB 316 TGTTCCTAAATTTGGGGTATGTTGGTGGGATGGGTCGCAACAGGTTGACACTCAC 375
QY 287 CGTGTGTATTTGCTAACGAGAGGCAATTCAATGGGCTACAAAGAACAAAGTTCAAA 346
DB 376 CGTGTGTATGCTCAATCGAGGGTATATCATGGGCAACAAAGAACAAATTCATAA 435
QY 347 AGCCAATTAATTTGGTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCAGAG 406
DB 436 AGCTAATCTATTTGGATTCCTCACTCAAGGCTTCTTATTCGTGTGTGTTTAAATGG 495
QY 407 AGAGAAATCTATGCCCCATTCAAGAGCTGTCTCAATAGTTAACTCCAGACATTTGT 466
DB 496 AGAGAGATTTATGCTCTTTTAAAGACCTTCTTCTTAATGTGAACCCAGATATATAGT 555
QY 467 GTTTGGGGATGAGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGCAAGGT 526
DB 556 GTTTGGGGGCTGGGACATAGTGAACCTGAGACGCGCATGACGAGGGCAAAAGGT 615
QY 527 GTTTGACATCGATTTTGCAGAGAGGTTGAGGCTTACATGGAATCCATGCTTCCATCCC 586
DB 616 GTTTGACATGATTTTGCAGAAAGAGGTTGAGACTTACATGAGACATAGTTCCCTTCC 675
QY 587 CGGAATCTATGACCCGGATTTCTATGCTGCAACCAAGAGAGCGTTCACCAACCTCAT 646
DB 676 TGGTATCTATGACCCGATTTCTATGCTGTAACCAAGATTCCTGCGAAATATACATCAT 735
QY 647 CAAGGGCAAAAGAGAGCAAGACTTCAAAATCATCAAGAACATGAAGCGCTTAAAGA 706
DB 736 TAAGGGCACTTAAGAGAGCAAGATTAAGCAATTAAGAGACATGAGGAATTTAAGA 795
QY 707 AGCCACCAAAAGTGAAGAGGTGTTGATCTGTGCACTGCCAACACAGAGAGTACAGTAA 766
DB 796 GAAGAACCAAGTGAAGAGGTGTTGATCTGTGCAAGCAACACAGAGAGTACAGTGA 855
QY 767 TTTGTTGTGGGCTTAAATGACCAATGAGAAATCTTTGGCTGCTGTGGACAGAAATGA 826
DB 856 TGTGTTGTGGGCTCAATGACCAACAGAGAACTTATAGGACGCTGTTGAAGAATGA 915
QY 827 GGTGAGATTTCTCTTCCACTTGTATGCAATGCTGTGTATGGAATAATGTTCTTT 886
DB 916 GTCTGAGATTTCTCAATGCACTTGTATGCTATGCTTGATGATCAAGAGAGATGCTCTTT 975

```

```

QY 887 CATTAATGAAGACCTTCAGAAACATTTTGTACAGAGGCTGATTTGATCTTGGCATGCGAG 946
DB 976 CATCAACCGTATGCTTCAGAACACCTTGTGCTGGTCTCATTTGAATCGGCGATAAATAA 1035
QY 947 GAACCTTTGATTTGGTGAAGATGAATCTTCAAGAGTGTGACGACCAAAATGAATCTGTGT 1006
DB 1036 GAATGTTTGAATGAGAGATGATTCAGAGAGTGTGACAGCTTAAGATGAATAATGATGCT 1095
QY 1007 GATTGATTTCTTGTGGGGGCTGTGATCAAGCAACATCTATAGTCAAGTTACMAACATCT 1066
DB 1096 CGTTACCTTCTTGTGTGGGGCTGTGATTAAGCAACATCAATTTGAGACTCAACACATCT 1155
QY 1067 GGGAAACATGATGATGATGAATCTTTGGCTTCACAACTTTCCGTTCCAGAGAAATCTC 1126
DB 1156 GGGAAACATGATGATGATGAACCTTTTCAGACCTTCAGACCTTCAGAGTTAAAGAAATTC 1215
QY 1127 CAAGAGCAACGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1186
DB 1216 AAGAGTAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1275
QY 1187 ACATCCAGACCATGTTGTTGTTATTAAGTATGATGATGATGATGATGATGATGATGATG 1246
DB 1276 ACACCTGATCATGTTGTTGATCATAGATGATGATGATGATGATGATGATGATGATGATG 1335
QY 1247 CATGATGATGATCACTTCAAGATATTTCAATGATGATGATGATGATGATGATGATGATG 1306
DB 1336 TATGATGATGATCAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1395
QY 1307 CACATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1366
DB 1396 CACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1455
QY 1367 GCTCAGACTGAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1426
DB 1456 GCTTGAACACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1515
QY 1427 AGTTCCTACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1486
DB 1516 AGTGGCTACATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1575
QY 1487 GGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1546
DB 1576 GGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1632
QY 1547 ATTGGCCGAGAGATTAACATGATCTCGAGTACAAAGTGAAG 1588
DB 1633 TTTAGCTCCAGAGAACACATGATCTTGAAGTACAAAGTGAAG 1674

```

RESULT 15
AY065415 1863 bp mRNA linear PLN 18-SEP-2002
LOCUS Arabidopsis thaliana putative myo-inositol-1-phosphate synthase
DEFINITION (At5g10170) mRNA, complete cds.
ACCESSION AY065415
VERSION AY065415.1 GI:17529257
KEYWORDS FLU CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1863)
REFERENCE
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banb, J., Chung, M.K.,
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,
Carlini, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishide, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis full length cDNA clones
TITLE
JOURNAL
REFERENCE
2 (bases 1 to 1863)

AUTHORS Yamada, K., Banh, J., Chang, C.H., Chang, B., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE Submitted (04-DEC-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PEGC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J., Chang, C.H., Chang, B., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PEGC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PEGC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES Location/Qualifiers

source 1. .1863
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAF09-36-B07 (R18590)"
/ecotype="Columbia"
/note="This clone is in a modified pBluescript vector (pLC-1) as a BamHI/XhoI insert."

gene 1. .1863
/gene="At5g10170"
1. .90
/gene="At5g10170"
91. .1623
/gene="At5g10170"
/codon_start=1
/evidence=experimental
/product="putative myo-inositol-1-phosphate synthase"
/protein_id="AAL38856.1"
/db_xref="GI:17529258"
/translation="MFISFQVSPNPKYKTEHSHVYDQTTTELHKNKGAPOFTV KRTVYEFKTDTHVPLKGLVGLVGMGNGSITLTAQVIANRGIISMAETKIQQAAYF GSLTQASSIRVGSFNGEETIYAPFKSLPNNPBEIYFGGNDISDMKLAAMAKYLD IDLQKMRPFHEHWPPLPGIFDPDITAAQSSRAHVIKTKQOLEQVIAKIREKE KKKVDVAVLVWANTERYSNVVGANDITENLMSLDKDEAEISPTLEIAICVLENY PFINSPOPTFWPGLIELAKIKNCLIGDGFSGTKMSVLDVFLVAGIKPTSLVS YNHLGNNDMNLAPQTPRSKEISKNVVDVSGNGLIYBGEHPDHVVIKYPCV GSKRAMDEYTSIEIPGKQKTIYMHNTCEDSILAAPILIDIVIALETTRIOPMSNE GKPHSFHVAITLISYLSKAPLVPPTGVVNALSKORAMLENVLRACVGAPENNMLE YK"

CDS 1624. .1863
/gene="At5g10170"
1846
/gene="At5g10170"
/note="compared to genomic sequence"
/replace="c"

3' UTR 1624. .1863
/gene="At5g10170"
1846
/gene="At5g10170"
/note="compared to genomic sequence"
/replace="c"

misc_difference 1846
/gene="At5g10170"
/note="compared to genomic sequence"
/replace="c"

ORIGIN

Query Match 57.4%; Score 1010.8; DB 8; Length 1863;
Best Local Similarity 78.3%; Pred. No. 1.7e-261;

Matches 1213; Conservative 0; Mismatches 337; Indels 0; Gaps 0;

QY 50 AATAATGTCATGAGAAATTTTAAGTTGAGTCTCTAATGTGAATGACCGAGACTGA 109
|||
DB 87 AAAAAAGTTCATGAAAGCTTCAAGGTGATCTCCGAACGTGAAGTACCCAGATGA 146
QY 110 GATTCAGTCGCTGTACAACTAGCAAAACCCGAACTGTTTACAGAGAAAGGAATGAC 149
|||
DB 147 GATTCAGTCGCTGTACAACTAGCAAAACCCGAACTGTTTACAGAGAAAAACGCTGC 206
QY 170 CTATCAGTGGATTTGTCAAAACCCAAATCTGTCAATATGAAATTTAAACCAATCCATGT 229
|||
DB 207 TTTTCAATGGAAGTGTGAAGCAAAAGCTGAAAGTATGAGTTTAAGACGAGCATATGT 266
QY 230 TCTTAATTAAGGGGTATGCTGTGTGGGTGGGGTGAACAACGGCTCAACCTCACCGG 289
|||
DB 267 TCCCAAACTTGGAGTAACTCTGTTGTTGGGAGGCAACAATGTCTTACACTTACCGC 326
QY 290 TGGTGTATTTGCTAACCGAGAGGCGATTTCAATGGCTTACAAAGGACAAGTTGAACAGC 349
|||
DB 327 TGGTGTATTTGCTAACCGAGAGGCGATTTCAATGGCTTACAAAGGACAAGTTGAACAGC 386
QY 350 CAATTACTTGGCTCCCTCACCAACCTCAGCTATCCGAGTTGGCTCTTCCAGGAGAG 409
|||
DB 387 TAATTAATTTCCGGTCAATTAAACCAACGATCTTATTCGGTGGGTCAATTAATGTGA 446
QY 410 GGAATCTATGCCCCATTCAGAGCCCTGCTTCCATATGTTAACTTGACGACATTTGTT 469
|||
DB 447 AGAAATCTATGCTCTTCAAAAGCTTCTCCCATGTGAATCCGGAAGAAATTTGTT 506
QY 470 TGGGAGATGGATATATGACAACTGAACCTGGCTATGCGATGCGCAAGGCAAAAGTGT 529
|||
DB 507 TGGAGATGGAGATTAAGTATGATGAATGAATTTAGCAGCAATGGCAAGGCAAGTCT 566
QY 530 TGACATGCAATTTGCAAGAACGATGAGGCTTACATGATGCAATGCTTCCATCCCGG 589
|||
DB 567 AGACATTCGACTTCAAAACCAATGATGCTGCTTCAATGAAACATGATGCTTCCATCCCGG 626
QY 590 AATCTATGACCCGATTTTCAATGCTGCGCAACCAAGAGAGCGTCCCAACAGCTCATCA 649
|||
DB 627 AATCTTGACCCGATTTTCAATGCTGCGCAACCAAGAGAGCGTCCCAACAGCTCATCA 686
QY 650 GGGCAACAAAGCAAGCAAGTCAACAATCATTCAAAGCATCAAGCGCTTTAAGGAAG 709
|||
DB 687 AGGCAACAAAGCAAGCAAGTCAACAATCATTCAAAGCATCAAGCGCTTTAAGGAAG 746
QY 710 CACCAAGTGAACAAGGTGTGTACTGTGACTGCGCAACAGAGAGGTACAGTAATT 769
|||
DB 747 AAACAAAGTGAACAAGGTGTGTACTGTGACTGCGCAACAGAGAGGTACAGTAATT 806
QY 770 GGTGTGGGCTTAAATGACACATGAGATCTTTGCTGCTGTGAGACAGAAATGAGG 829
|||
DB 807 GGTGTGTGCTTAAATGACACATGAGATCTTTGCTGCTGTGAGACAGAAATGAGG 866
QY 830 TGAATTTTCTTCTCACTTGTATGCAATGCTGCTGTGTATGAGAAATGCTTCTTCA 889
|||
DB 867 TGAATTTTCTTCTCACTTGTATGCAATGCTGCTGTGTGTGAGAAATGCTTCTTCA 926
QY 890 TATGGAAGCCTCAAGAACCTTTGTATGCAAGGCTGATGTAATTTGCAATGCGAGAA 949
|||
DB 927 CAATGGAAGTCAAGAACCTTTGTATGCAAGGCTGATGTAATTTGCAATGCGAGAA 986
QY 950 CAATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1009
|||
DB 987 CTGCTTGAATCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1046
QY 1010 TGAATTTTCTTGTGGGGGCTGTGATGCAAGCAATCTATGATGATGATGATGATGATG 1069
|||
DB 1047 CGATTTCTTGTGTGGGGGCTGTGATGCAAGCAATCTATGATGATGATGATGATGATG 1106
QY 1070 AAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129
|||
DB 1107 GAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1166

```
QY 1130 GAGCAACGTTGTTGATGATATGATGCAACAGCAATGCCATCTCTATGAGCCTGTGAACA 1189
    |||||
Db 1167 GAGCAATGTCGTGAGACGATATGTTGTGACAAACGAAATTTGTAGACCTGTGAACA 1226
    |||||
QY 1190 TCCAGACCAATGTTGTTGTTATTAAAGTATGTCCTTAAGTGGGACAGCAAGAGCCAT 1249
    |||||
Db 1227 CCTGACCAAGTTGTGATCATTAAGTATGTGCATGTGTAGAGACAGCAAGAGCAAT 1286
    |||||
QY 1250 GGATGATTCATTTGAGAGATATTCATGGGTGAAAAGACCATTTTTCACACAC 1309
    |||||
Db 1287 GGATGATTAACATACATAGAGATATTCATGGGAGGAAACACGATTTGATGCACACAC 1346
    |||||
QY 1310 ATGCGAGATTCCTCTTAGCTGCTGCTATTAATCTTGACTGTGCTCTTCTGTAGCT 1369
    |||||
Db 1347 TTGTGAAGACTCTCTCTTGGCTGCTCAATTAATTTGGACTTGTCTTCTTCTGTAGCT 1406
    |||||
QY 1370 CAGCACTAGATCGAGTTTAAAGCTGAAAATGAGGAAAATTCACACTCATTCACCCAGT 1429
    |||||
Db 1407 CACTACTGGATCCAGTTCATGAGTGAAAACGAGGAAAGTTTCATTTCCATCCGCT 1466
    |||||
QY 1430 TGGTACCATCTCAGCTACCTCACCAAGCTCTGTGTTCCACCGGTACACCACTGCT 1489
    |||||
Db 1467 GGCCTACTCTCTTAAGTTACTCTCAAAAGGCACCACTGCTTCTCTGACACGCCAGTTGT 1526
    |||||
QY 1490 GAATGATTTGTCAAAGCAGCGTCAATGCTGAAAACATATAGAGGGCTTGTGTGATT 1549
    |||||
Db 1527 TAATGCTTTGTGAAACAGCGTCAATGCTTGAAGACGTACTCAGAGCTTGTGTGACT 1586
    |||||
QY 1550 GGCCCCAGAGAAATTAACATGATTTCTCGAGTACAAAGTAGACATGGGACCGA 1599
    |||||
Db 1587 TGCTCTGAGAACAAATCATGATTTCTCGAGTACAAAGTAGATCAAGAGACCAA 1636
    |||||
```

Search completed: June 8, 2005, 02:53:56
Job time : 7692.08 secs